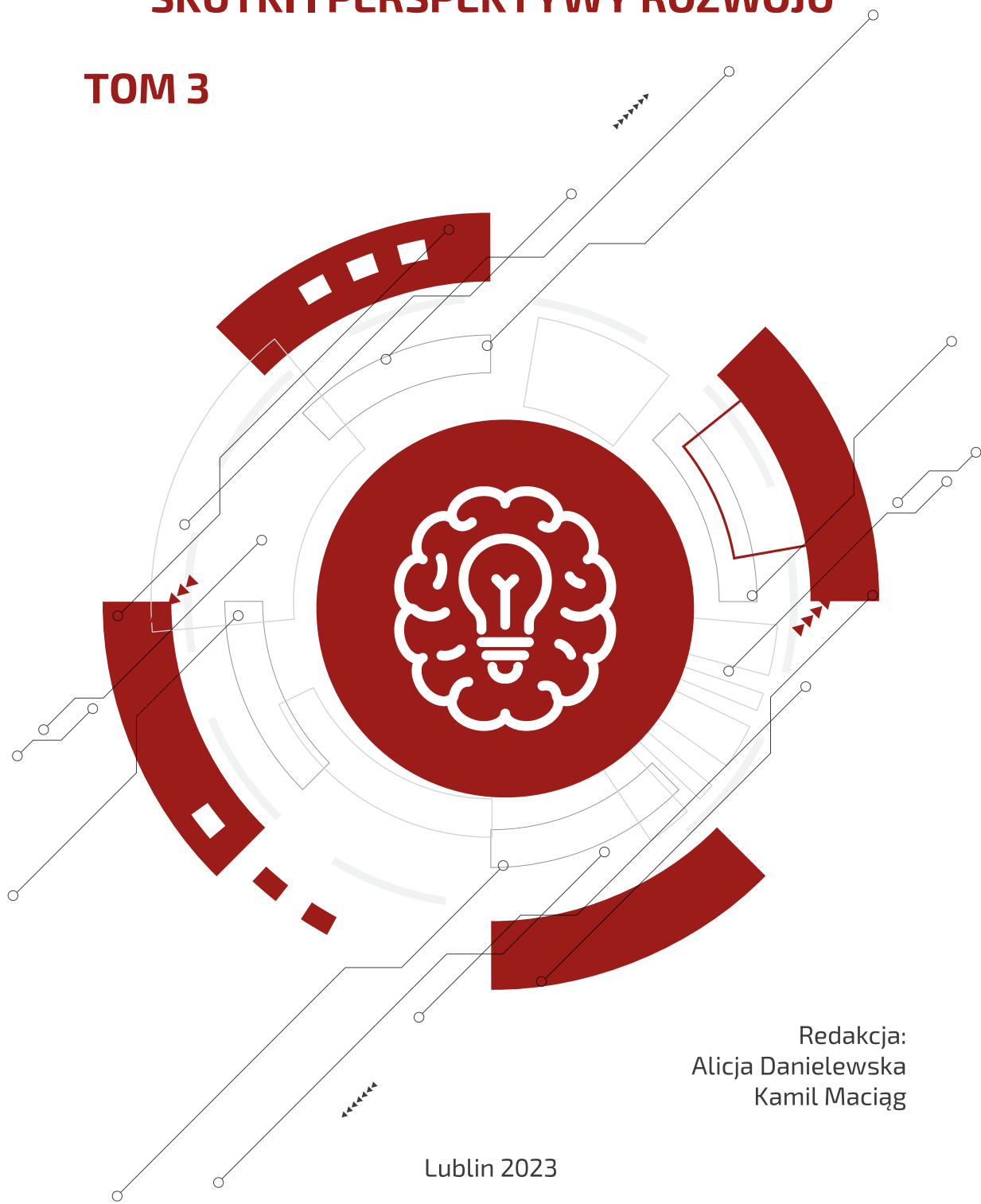


ROZWIĄZANIA TECHNOLOGICZNE XXI WIEKU SKUTKI I PERSPEKTYWY ROZWOJU

TOM 3



Redakcja:
Alicja Danielewska
Kamil Maciąg

Lublin 2023

**Rozwiązania technologiczne
XXI wieku –
skutki i perspektywy rozwoju**

Tom 3

**Rozwiązania technologiczne
XXI wieku –
skutki i perspektywy rozwoju**

Tom 3

Redakcja:
Alicja Danielewska
Kamil Maciąg

Lublin 2023

**Wydawnictwo Naukowe TYGIEL składa serdeczne podziękowania
zespółowi Recenzentów za zaangażowanie w dokonane recenzje
oraz merytoryczne wskazówki dla Autorów.**

Recenzentami niniejszej monografii byli:

- prof. dr hab. Katarzyna Cieślak-Blinowska
- prof. dr hab. inż. Zbigniew Konopka
- dr hab. Agnieszka Andrychowicz-Trojanowska
- dr hab. inż. Mirosław Bramowicz, prof. UWM
- dr hab. Łukasz Łaczmański, prof. IITD PAN
- dr hab. Piotr Piszczek, prof. UMK
- dr inż. Jarosław Bednarz
- dr n. o zdr. Mariola Janiszewska
- dr Arkadiusz Marcin Rudzki
- dr Andrzej Wartecki

Wszystkie opublikowane rozdziały otrzymały pozytywne recenzje.

Skład i łamanie:
Monika Maciąg

Projekt okładki:
Marcin Szklarczyk

Korekta:
Ewelina Chodźko

© Copyright by Wydawnictwo Naukowe TYGIEL sp. z o.o.

ISBN 978-83-67104-75-3

Wydawca:
Wydawnictwo Naukowe TYGIEL sp. z o.o.
ul. Głowackiego 35/341, 20-060 Lublin
www.wydawnictwo-tygiel.pl

Spis treści

Stanisław Andrzej Różański

Ciekłe kryształy domieszkowane nanocząstkami i ograniczone geometrycznie – nowe zjawiska fizyczne i możliwości technologiczne 7

Justyna Czajka

Charakterystyka spektroskopowa oraz właściwości aplikacyjne nowych up-konwersyjnych nanomateriałów nieorganicznych 26

Joanna Malarz, Katarzyna Leszczyńska-Sejda, Grzegorz Benke, Karolina Goc, Karolina Pianowska, Julita Sztandera, Michał Ochmański

Heterogeniczne katalizatory renowe 35

Waldemar Łasica

Wpływ dodatku odpadowych wiórów stalowych na wartości modułu Young’ a eko-kompozytów cementowo-szklanych 52

Aleksander Mądrała, Michał Lentner

Stanowisko spawalnicze ramy elektrycznego motocykla LEM Photon 76

Radosław Belka

Niedeterministyczny algorytm zachłanny do zastosowań w strategii marszrutyzacji – studium przypadku 93

Agnieszka Bier, Zdzisław Sroczyński

Głosowe interfejsy użytkownika: zastosowania, projektowanie, zagrożenia 114

Vladimir Golovko, Marta Chodyka, Piotr Lichograj, Konrad Grzegórski

Zastosowania i nowe rozwiązania dla głębokich sieci neuronowych 124

Piotr Lichograj, Marta Chodyka, Vladimir Golovko, Tomasz Bujnik

Metoda wykrywania napadów padaczkowych za pomocą fal EEG 134

Karolina Gabor-Siatkowska, Artur Janicki

Badanie zasadności używania kalibracji w pomiarach okulograficznych w celu zastosowania w interfejsie człowiek-komputer 148

Indeks Autorów 160

Ciekle kryształy domieszkowane nanocząstkami i ograniczone geometrycznie – nowe zjawiska fizyczne i możliwości technologiczne

1. Wprowadzenie

Ciekle kryształy (CK) są klasą niezwykłych materiałów, które znalazły szerokie zastosowanie w technikach dotyczących obrazowania informacji oraz w fotonice [1-4]. Wykorzystuje się je w ekranach telewizyjnych, monitorach komputerów, różnego rodzaju wyświetlaczach (np. w zegarkach) oraz innym sprzęcie codziennego użytku. CK są substancjami organicznymi o dość złożonej strukturze molekularnej. Nowe materiały ciekłokrystaliczne otrzymuje się głównie na drodze syntezy chemicznej. Jednak w ostatnich latach ta dziedzina badań zmienia się dynamicznie w kierunku nanonauki i nanotechnologii [5]. Nastąpił gwałtowny rozwój metod związanych z modyfikacją materiałów za pomocą różnego rodzaju domieszek, które stanowią, np. funkcjonalne nanocząstki ferromagnetyczne i ferroelektryczne [6]. Interesujące własności fotoniczne uzyskuje się poprzez domieszkowanie CK nanocząstkami krzemionki płomieniowej (aerosil), nanocząstkami złota (GNP, ang. *Gold Nano-Particles*), kropkami kwantowymi (QD, ang. *Quantum Dots*), cząstkami metali, półprzewodników, nanorurkami grafenowymi itp. [7].

Wprowadzenie nanocząstek do CK umożliwia modyfikację ich własności fizycznych lub chemicznych bez potrzeby syntezy nowych materiałów [8]. Obecność nanocząstek o odpowiedniej koncentracji wpływa na modyfikację własności elektrooptycznych, przewodnictwa czy zwiększenie stabilności określonych mezofaz, np. fazy niebieskiej (BP, ang. *Blue Phase*) [9]. Znaczej zmianie ulegają własności termodynamiczne, takie jak charakter i temperatura przejścia fazowego, a także dynamika relaksacji procesów molekularnych i kolektywnych [10, 11].

Interesującym aspektem łączenia własności CK i nanocząstek jest wykorzystanie samoorganizacji molekuł jako szablonu do odpowiedniego porządkowania i orientowania nanocząstek [12]. Ponadto anizodiametryczne nanocząstki o kształcie nanoprętów, nanorurek czy płatków mogą formować, przy odpowiednim stężeniu rozpuszczalnika, fazy charakterystyczne dla liotropowych ciekłych kryształów (LCK) [13].

Niektóre z omówionych powyżej efektów związanych z domieszkowaniem za pomocą nanocząstek przejawiają się również w przypadku CK ograniczonych geometrycznie, np. po umieszczeniu w membranach porowatych, szklach porowatych, itp. [14]. W zależności od struktury matrycy oraz rozmiarów porów główną rolę odgrywają oddziaływania z powierzchnią, czy efekty rozmiarów skończonych [15, 16]. Zetknięcie się CK z nanomateriałami prowadzi do uzyskania dodatkowych funkcjonalności, które różnią się od własności danej mezofazy oraz zastosowanych nanocząstek.

¹ srozansk@asta-net.com.pl, Akademia Nauk Stosowanych im. Stanisława Staszica w Pile, www.ans.pila.pl.

W niniejszej pracy starano się odpowiedzieć na następujące pytania:

- Jak modyfikować znane materiały ciekłokrystaliczne, aby udoskonalić ich własności fizykochemiczne?
- Jaki wpływ mają rozmiary, kształt i własności powierzchni nanocząstek na modyfikację materiałów oraz otoczenia fizycznego, w którym się znajdują?
- Jakie mogą być zastosowania praktyczne nanomateriałów ciekłokrystalicznych?

Ponadto dokonano przeglądu najnowszych osiągnięć w dziedzinie materiałów ciekłokrystalicznych domieszkowanych nanocząstkami oraz wpływu ograniczeń geometrycznych na własności fizyczne różnych mezofaz.

2. Ciekłe kryształy

Ciekłe kryształy to materiały anizotropowe charakteryzujące się zarówno właściwościami cieczy, jak i ciała stałego, a wiele parametrów fizycznych jest pośrednich między cieczą i kryształem. W zależności od sposobu otrzymywania fazy ciekłokrystalicznej wyróżniamy dwa rodzaje CK [17, 18]:

- termotropowe ciekłe kryształy – otrzymuje się przez ogrzanie ciała krystalicznego lub schłodzenie cieczy izotropowej, a kolejne przejścia fazowe zachodzą w wyniku zmiany temperatury;
- liotropowe ciekłe kryształy – powstają w wyniku rozpuszczenia w rozpuszczalniku izotropowym związków amfifilowych, polipeptydów, itp. W tym przypadku zmiana stężenia jednego ze składników powoduje powstanie odpowiedniej fazy ciekłokrystalicznej.

2.1. Termotropowe ciekłe kryształy

Termotropowe ciekłe kryształy, zgodnie z nomenklaturą wprowadzoną przez Friedela [19], można podzielić ogólnie na trzy klasy: nematyki, cholesteryki i smektyki. Różnią się one własnościami optycznymi oraz przestrzennym uporządkowaniem molekuł. Można je również podzielić ze względu na kształt molekuł, z których są zbudowane. Kształt cząsteczek może być np. prętopodobny i walcowy (kalamityczne CK), dyskopodobny (dyskotyczne CK) oraz bananowy (fazy CK tworzone przez molekuły niechiralne w kształcie banana).

2.1.1. Nematyczne ciekłe kryształy

W ciekłokrystalicznej fazie nematycznej (NCK) nie występuje długozasięgowe uporządkowanie środków mas molekuł, natomiast ich długie osie zorientowane są średnio wzdłuż wyróżnionego kierunku, który wyznacza wektor jednostkowy \mathbf{n} zwany direktorem. Nematyki są kryształami optycznie jednoosiowymi, z osią optyczną równoległą do kierunku direktora. Niemal we wszystkich przypadkach występuje symetria względem obrotu wokół osi direktora, a stany direktora $+\mathbf{n}$ i $-\mathbf{n}$ są nierozróżnialne. Ze względu na własności symetrii nematykowi przypisuje się punktową grupę symetrii $D_{\infty h}$ według oznaczeń Schoenfliesa [20].

2.1.2. Cholesteryczne ciekłe kryształy

Cholesteryczne CK (Ch) lokalnie nie różnią się uporządkowaniem od fazy nematycznej. Jednak w odróżnieniu od nematyków direktor \mathbf{n} zmienia swoje położenie w przestrzeni, tworząc strukturę helikoidalną, której symetria odpowiada grupie

punktowej D_{∞} . Ponieważ stany direktora $+n$ i $-n$ są nierozróżnialne, to okres struktury l w cholesteryku jest równy połowie nominalnego skoku helisy p ($l = p/2$). Budowa przestrzenna cholesteryków zbudowanych z asymetrycznych molekuł powoduje, że mają one unikalne własności optyczne, takie jak wyjątkowo duża aktywność optyczna oraz wykazują selektywne odbicie światła.

Podczas przejścia z fazy izotropowej do helikalnej fazy cholesterycznej, w bardzo wąskim zakresie temperatury, pojawia się silnie uporządkowana faza ciekłokrystaliczna zwana fazą niebieską. BP posiada interesujące własności aplikacyjne z punktu widzenia nowej generacji wyświetlaczy, a mianowicie bardzo krótki czas przełączania oraz szeroki kąt widzenia.

2.1.3. Smektyczne ciekłe kryształy

W smektycznych CK molekuły ułożone są w warstwach. Wyróżniamy dużą różnorodność faz smektycznych w zależności od rodzaju molekuł oraz ich uporządkowania i orientacji w poszczególnych warstwach:

- faza smektyczna A (SmA) – molekuły ustawione są długimi osiami średnio prostopadle do płaszczyzn warstw smektycznych, a ich środki mas rozłożone są nieregularnie w tej płaszczyźnie; każda warstwa może być traktowana jako ciecz dwuwymiarowa; faza SmA jest optycznie jednoosiowa o symetrii $D_{\infty h}$;
- faza smektyczna B (SmB) – środki mas molekuł w tej fazie ustawione są heksagonalnie, a ich warstwy można traktować jak dwuwymiarowe ciało stałe;
- faza smektyczna C (SmC) – długie osie molekuł w tej fazie nachylone są pod pewnym kątem do warstw smektycznych, tworząc strukturę dwuosiową optycznie o symetrii opisanej grupą punktową C_{2h} ;
- faza chiralna C (SmC*) – fazę tę tworzą molekuły chiralne, które wykazują aktywność optyczną; prowadzi to do powstania struktury helikoidalnej, w której przy przejściu od warstwy do warstwy uporządkowanie długich osi molekuł zmienia orientację w przestrzeni (grupa punktowa C_2); chiralne smektyki C wykazują własności ferroelektryczne wykorzystywane w urządzeniach elektrooptycznych.

Oprócz uporządkowania ferroelektrycznego wyróżnia się fazy o uporządkowaniu antyferroelektrycznym (SmC_A*), ferrielektrycznym (SmC_γ*) oraz wiele innych faz pośrednich [21]. Znanych jest również szereg innych, bardziej egzotycznych faz ciekłokrystalicznych, np. typu E, G, H, itp. [22].

2.2. Liotropowe ciekłe kryształy

Liotropowe CK otrzymuje się ze zmieszania dwóch lub więcej składników. LCK można podzielić na trzy różne rodzaje: amfifilowe, koloidalne oraz chromoniki, gdzie w tych ostatnich cząsteczki składowe są cząsteczkami barwnika w odpowiednim rozpuszczalniku. Powstają one zwykle podczas rozpuszczania molekuł amfifilowych (np. detergentów, mydeł, itp.) w wodzie. LCK mają duże znaczenie biologiczne, gdyż mogą być zbudowane z więcej niż dwóch składników (np. układy lipid-woda, sole kwasów trawiennych w wodzie, fosfolipidy-woda, lipid-woda-białko) przy czym woda jest integralną częścią tych struktur. W zależności od stężenia rozpuszczalnika oraz temperatury formowane są różne struktury związane z geometrycznym rozkładem molekuł amfifilowych, takie jak struktura lamelarna (warstwowa), kubiczna, walcowa heksagonalna, micelarna i na koniec roztwór jednorodny [23].

3. Nanomateriały

Nonomateriały to materiały nanostrukturalne, których rozmiary w co najmniej jednym z kierunków mieszczą się w zakresie od 1 nm do 100 nm [24]. Posiadają zależne od rozmiarów własności elektryczne, magnetyczne, optyczne i chemiczne, które znacznie różnią się od własności materiałów w postaci objętościowej, a także pojedynczych cząsteczek lub atomów.

Nanocząstki charakteryzuje się pod względem fizycznym i chemicznym za pomocą następujących metod i procedur [25]:

- mikroskopia sił atomowych – obrazowanie nanocząstek;
- mikroskopia elektronowa – obrazowanie nanocząstek;
- dynamiczne rozpraszanie światła laserowego – pomiar rozkładu wielkości nanocząstek;
- dyfrakcja promieniowania rentgenowskiego – pomiar rozmiarów i rozkładu wielkości nanokryształów;
- spektroskopia w podczerwieni z transformatą Fouriera;
- spektroskopia elektronowego rezonansu paramagnetycznego.

Wyróżnia się cztery podstawowe grupy nanomateriałów, które wykorzystuje się jako domieszki do ciekłych kryształów [26, 27]:

- nanocząstki metali, np. złota, srebra, miedzi, żelaza;
- nanocząstki ferroelektryczne, np. tlenek barowo-tytanowy BaTiO_3 ;
- nanocząstki półprzewodnikowe, np. nanodruity i kropki kwantowe – CdS , CdSe , CdTe , InP , InAs ;
- nanocząstki węgla – nanorurki, fulereny, grafen, włókna węglowe.

3.1. Nanocząstki w termotropowych ciekłych kryształach

Bardzo ciekawą klasę nowych nanomateriałów ciekłokrystalicznych stanowią ferromagnetyki. Są to zawiesiny nanocząstek magnetycznych (np. nanocząstek magnetytu, nanorurek węglowych sfunkcjonalizowanych cząstkami ferromagnetycznymi, itp.) w NCK, przygotowane w celu zwiększenia czułości układu na pole magnetyczne [28]. Teoria zawiesin magnetycznych została podana w latach siedemdziesiątych ubiegłego wieku i od tego czasu pojawiło się bardzo dużo wyników badań tego typu materiałów [29]. Domieszkowanie nematyków nanocząstkami magnetytu o różnych kształtach (kuliste, prętowe, łańcuchowe) prowadzi do wniosków, że występuje słabe kotwiczenie molekuł NCK na cząstkach kulistych, natomiast w przypadku cząstek prętowych i łańcuchowych to kotwiczenie jest silne, z równoległą początkową orientacją pomiędzy momentami magnetycznymi cząstek magnetycznych i директором [30]. Okazuje się, że dodanie nanocząstek magnetycznych do CK obniża temperaturę przejścia fazowego. Co ciekawe, efekt zależy od kształtu nanocząstki i jest silniejszy dla nanocząstek w postaci pręta w porównaniu do nanocząstek sferycznych. Obliczona energia aktywacji zmienia się niemonotonicznie i jest również zależna od kształtu domieszek magnetycznych [31]. Innym ciekawym przykładem jest układ NCK domieszkowanego wielościennymi nanorurkami węglowymi, w którym następuje przesunięcie napięcia progowego Freédericksza w stronę mniejszych wartości wraz z dodatkiem domieszek [32].

Nową możliwością zastosowania nanokompozytów jest wykorzystanie CK domieszkowanych nanocząstkami ferroelektrycznymi, np. typu tlenku barowo-tytanowego

BaTiO₃ do zapisywania i przechowywania informacji [33]. Obecność nanocząstek BaTiO₃ modyfikuje własności elektrooptyczne mezofazy bez zmiany uporządkowania nematycznego. Efekt pamięci trwałej został odkryty w nanokompozycie utworzonym przez 5CB (5CB, ang. *4-n-pentyl-4'-cyanobiphenyl*) i BaTiO₃ [34]. W fazie izotropowej 5CB, w pobliżu nanocząstek BaTiO₃ tworzą się domeny pseudonematyczne rozłożone statystycznie w fazie izotropowej. Po przyłożeniu pola elektrycznego domeny pseudonematyczne ulegają reorientacji w kierunku pola elektrycznego. Po wyłączeniu pola domeny pozostają zorientowane – obserwujemy efekt pamięci. Wielkość histerezy dielektrycznej układu 5CB-BaTiO₃ zależy od stężenia nanocząstek BaTiO₃ w fazie izotropowej 5CB. Udekorowanie powierzchni nanocząstek BaTiO₃ polimerami blokowymi prowadzi do gigantycznego efektu pamięci, który objawia się występowaniem histerezy dielektrycznej związanej również z powstaniem domen pseudonematycznych w badanym kompozycie [33].

Efekt pamięci zaobserwowano także w nanokompozytach ciekłokrystalicznych złożonych z NCK i aerosilu (FN, ang. *Filled Nematics*). Efekt pamięci zależy w znacznym stopniu od koncentracji nanocząstek aerosilu w nematyku, od stanu powierzchni nanocząstek (powierzchnia hydrofilowa lub hydrofobowa), oddziaływań molekularnych nematyk-aerosil oraz parametrów przyłożonego pola elektrycznego [35-37]. Ponadto, po przyłożeniu pola elektrycznego do uprzednio mętnej komórki zawierającej FN staje się ona przezroczysta. Komórka zachowuje przezroczystość po wyłączeniu pola, co wynika z faktu, że powierzchnia nanocząstek aerosilu jest około dwa rzędy wielkości większa od powierzchni elektrod. Powrót do stanu rozpraszającego światło jest możliwy, np. przez ogrzanie próbki nieznacznie powyżej punktu klarowania [37]. W celu uzyskania efektu pamięci wykorzystano także układy koloidalne bazujące na CK domieszkowanym cząstkami niemagnetycznymi (np. SiO₂) oraz cząstkami magnetycznymi (np. Fe₃O₄). Efekt pamięci objawia się histerezą dielektryczną fazy nematycznej w temperaturze pokojowej. Obecność nanocząstek magnetycznych umożliwia zmianę wielkości efektu pamięci po przyłożeniu pola magnetycznego [38].

Zaobserwowano, że dynamika procesów molekularnych związanych z rotacją molekuly wokół krótkiej osi (proces δ) oraz modu libracyjnego związanego z oscylacjami molekuly wokół kierunku direktora w NCK 5CB ulega zmianie wraz ze wzrostem koncentracji aerosilu hydrofilowego i hydrofobowego. Odkryte prawidłowości w przesunięciach częstotliwości procesów relaksacji powiązано ze zmianami parametru porządku. Ponadto, w układach dyspersyjnych 5CB-aerosil wykazano wpływ koncentracji aerosilu hydrofilowego i hydrofobowego na formowanie uporządkowania prenematycznego w fazie izotropowej, w pobliżu przejścia fazowego z fazy izotropowej do fazy nematycznej [10].

Ferroelektryczne CK (FCK) stanowią następną klasę materiałów, która jest intensywnie badana z punktu widzenia zastosowań fonicznych oraz w wyświetlaczach ciekłokrystalicznych, ze względu na niskie napięcie przełączania, krótki czas odpowiedzi elektrooptycznej, dobry kontrast optyczny, efekt pamięci, itp. [21]. Z dotychczasowych badań wynika, że domieszkowanie FCK nanocząstkami metali, tlenków metali, ferroelektryków, izolatorów, grafenu, półprzewodników, itp. prowadzi do znacznego polepszenia ich własności elektrooptycznych.

W FCK obserwuje się dwa specyficzne procesy relaksacji w obszarze niskich częstotliwości: mod Goldstone'a (GM) i mod miękki (SM). GM związany jest z fluktu-

acjami kąta azymutalnego, tzn. ustawieniem płaszczyzny pochylenia molekuł w warstwach smektycznych i obserwowany jest jedynie w fazie ferroelektrycznej. SM związany jest z fluktuacjami amplitudy parametru porządku (wartością kąta pochylenia molekuł w stosunku do normalnej do warstw smektycznych) i występuje zarówno w fazie ferroelektrycznej, jak i paraelektrycznej [21]. W układach dyspersyjnych FCK-aerosil hydrofilowy obserwuje się zmniejszenie inkrementu dielektrycznego GM i SM wraz ze wzrostem koncentracji aerosilu w mieszaninie, znikanie GM dla określonej koncentracji nanocząstek aerosilu, wpływ koncentracji aerosilu na częstotliwość modów oraz usunięcie degeneracji charakterystycznej częstotliwości GM i SM w przejściu fazowym SmA–SmC* wraz ze wzrostem koncentracji aerosilu [11]. Zaobserwowane prawidłowości wyjaśniono, biorąc pod uwagę deformację warstw smektycznych, rozbitcie fazy ferroelektrycznej na domeny oraz wprowadzenie do fazy ciekłokrystalicznej zamrożonego przypadkowego nieporządku w wyniku dodania do FCK nanocząstek aerosilu.

Szczególne zainteresowanie wzbudzają nanomateriały półprzewodnikowe, które prowadzą do poprawienia własności dielektrycznych, zwiększenia efektu pamięci, mocy, szybszego przełączania oraz poprawienia efektywności fotoluminescencji [39]. W zależności od rozmiarów ograniczeń geometrycznych dla ruchu elektronów wyróżnia się takie nanoobiekty, jak QD, pręty i druty kwantowe oraz studnie kwantowe. Z drugiej strony należy zwrócić uwagę, że wprowadzenie nanocząstek do CK może prowadzić do zaburzenia struktury oraz zmiany uporządkowania geometrycznego mezofazy, pogarszając parametry dielektryczne i elektrooptyczne. Niezbędna jest optymalizacja koncentracji nanocząstek w kompozycie ciekłokrystalicznym.

Okazuje się, że grupy nanocząstek w odpowiednich warunkach mogą wykazywać zachowanie kolektywne prowadzące do modyfikacji ich własności w porównaniu do próbek objętościowych oraz pojedynczych cząstek. Prowadzi to do nowych efektów wpływających na własności elektryczne, magnetyczne czy fotoniczne [40], które wynikają z oddziaływania między nanocząstkami. Szczególnie interesujące są układy QD rozproszonych w CK. Przejście z fazy izotropowej do fazy nematycznej wykorzystuje się do tworzenia zorganizowanych grup QD typu CdSe/ZnS [41]. W kontrolowanych warunkach dobrze rozproszone QD są wypychane z domen o uporządkowaniu nematycznym do współistniejących obszarów fazy izotropowej. W efekcie końcowym otrzymuje się uporządkowaną, trójwymiarową grupę QD ulokowanych w punktach defektów mezofazy. Ruch pojedynczych nanocząstek oraz ich zespołów można śledzić, wykorzystując fotoluminescencję QD.

Przykładem praktycznego wykorzystania kompozytów ciekłokrystalicznych na bazie QD jest laser kompozytowy, w którym zastosowano QD rozproszone w ChCK z dodatkiem molekuł chiralno-azobenzenowych [42]. Przestrzalność lasera zapewniona jest poprzez skracanie lub wydłużanie skoku helisy cholesterycznego CK indukowane promieniowaniem ultrafioletowym oraz naświetlaniem światłem niebieskim powodującym przejście *trans-cis* oraz *cis-trans* chiralnego azobenzenu. Ponadto omawiany laser charakteryzuje się dużą odpornością na zniszczenie ($>85 \mu\text{J}$ w impulsie) oraz dużą stabilnością optyczną.

Na szczególną uwagę zasługują układy dyspersyjne składające się z nanocząstek złota oraz termotropowych CK [43]. GNP wykazują zależne od rozmiarów własności elektryczne, optyczne i chemiczne wynikające z wymiarowych efektów kwantowych wpływających na odmienne zachowanie się elektronów w porównaniu do próbek

makroskopowych. Kolektywne oscylacje swobodnych elektronów w wyniku silnego oddziaływania fali elektromagnetycznej nazwano powierzchniowym rezonansem plazmonowym (SPR, ang. *Surface Plasmon Resonance*). Zależy on od rozmiarów, kształtu i stałej dielektrycznej GNP oraz od otaczającego medium, np. woda, alkohol, itp. Własności nanokompozytów można kontrolować poprzez dobranie odpowiedniego stężenia GNP. Domieszkowanie fazy nematycznej za pomocą GNP wpływa na przewodnictwo, temperaturę przejścia fazowego nematyk-faza izotropowa, napięcie progowe przejścia Fréedericksza, anizotropię dielektryczną, itp. Domieszkowane GNP smektycznej fazy ferroelektrycznej powoduje, jak już wspomniano powyżej, polepszenie własności elektrooptycznych, modyfikację dynamiki modów kolektywnych w fazie SmC* (efekt zależy również od rozmiarów GNP) [44], skrócenie czasów przełączania oraz zmniejszanie lepkości rotacyjnej.

W tym kontekście nie można pominąć bardzo istotnego zjawiska związanego ze stabilizacją fazy niebieskiej w cholesterycznych CK z wykorzystaniem GNP. W pracy [45] pokazano, że dodanie GNP do BP poszerza przedział temperatury istnienia fazy od 0,5 deg do 5 deg. Efekt wyjaśniono akumulacją GNP na liniach dysklinacji w BP. Jest to podobny mechanizm do stabilizowania BP za pomocą polimerów [46]. Ponadto wykazano, że zdyspergowanie nanoprętów złota w BP prowadzi do tego, że zakres tej fazy osiąga nasycenie przy krytycznym stężeniu nanoprętów. Własności optyczne kompozytu można kontrolować za pomocą pola elektrycznego i temperatury [47].

Innymi szeroko wykorzystywanymi domieszkami służącym do modyfikacji własności fizycznych CK są nanomateriały węglowe [48]. Należy tutaj wymienić między innymi fulereny, nanorurki węglowe (CNT, ang. *Carbon Nanotube*), grafen i tlenki grafenu, które wykazują niespotykane własności fizykochemiczne. CNT są zwiniętymi arkuszami grafenu w nanorurkę. Występują nanorurki jednościenne i wielościenne o długości około 1-100 nm i średnicy rzędu 1-4 nm. Oddziaływanie między molekułami CK i CNT prowadzi do formowania uporządkowanej struktury supramolekularnej. Tego typu struktury definiują unikalne własności fizyczne zawiesiny CK-CNT, takie jak przewodnictwo elektryczne, przenikalność elektryczna, przejścia fazowe, transmisja optyczna, efekt pamięci, które mogą znaleźć zastosowanie w elektrooptyce i optoelektronice [49]. Umieszczenie CNT w CK powoduje nieznaczne obniżenie temperatury przejścia fazowego, modyfikację lepkości, stałych sprężystości, itp. Obserwuje się silne oddziaływanie molekuł CK z powierzchnią CNT prowadzące do powstania mikronowych warstw ciekłokrystalicznych ze złożoną strukturą dwójłomną. Powierzchniowa chiralność CNT może być przenoszona na niechiralną fazę ciekłokrystaliczną. Zaobserwowano również wpływ pola magnetycznego oraz elektrycznego na reorientację CNT w fazie ciekłokrystalicznej. Ponadto, przewodnictwo elektryczne w układach CK-CNT zależy w sposób złożony od koncentracji CNT w CK, danej fazy ciekłokrystalicznej oraz temperatury. Obecność CNT o niewielkim stężeniu w antyferroelektrycznym CK wywiera znaczący wpływ na sekwencję wszystkich chiralnych mezofaz typu C oprócz zwykłej fazy SmC*.

Fulereny zbudowane z 60-70 atomów węgla (C₆₀, C₇₀) są sferycznymi molekułami o rozmiarach nanometrycznych. W celu polepszenia własności elektrycznych i elektrooptycznych CK dąży się do przyłączenia C₆₀ wiązaniem kowalencyjnym do molekuł CK, jak również domieszkuje się CK za pomocą C₆₀ [50]. Inną odmianą alotropową węgla jest grafen, który tworzy dwuwymiarowa monowarstwa atomów węgla uporząd-

kowanych heksagonalnie. Grafen jest słabo rozpuszczalny w CK, co ogranicza jego wykorzystanie. Natomiast tlenek grafenu dodany do CK sprzyja polepszeniu własności elektrooptycznych w wyniku sprzężenia dipoli elektrycznych na powierzchni tlenku grafenu z momentami dipolowymi molekuł CK, przyczyniając się do pułapkowania zanieczyszczeń jonowych.

3.2. Samoorganizacja nanocząstek

Koloidalne zawiesiny cząstek anizotropowych, w tym pręty, nanorurki, dyski, elastyczne łańcuchy i nanodruty mogą samoorganizować się od fazy izotropowej do fazy uporządkowanej nematycznie, gdy stężenie rozpuszczalnika przekracza wartość krytyczną zgodnie z modelem Onsagera [51-53].

Uporządkowania ciekłokrystalicznego CNT można oczekiwać dla nanocząstek o odpowiedniej długości. Teoria Onsagera przewiduje początek formowania fazy nematycznej dla ułamka objętościowego $v = 3,3d/x$, gdzie d oznacza średnicę nanorurki, a x jej długość. Typowy stosunek d/x dla jednościankowych CNT jest rzędu 10^{-3} . Wiele faz ciekłokrystalicznych tego typu zostało opisanych w literaturze [54]. Uporządkowanie ciekłokrystaliczne można również uzyskać dzięki chemicznej funkcjonalizacji powierzchni CNT lub wykorzystać do tego celu odpowiednie surfaktanty lub polimery. Przykładowo nanocząstki tlenku wanadu V_2O_5 w kształcie igiełek umieszczone w odpowiednim rozpuszczalniku mogą formować liotropową fazę nematyczną, której własności zależą od stężenia, temperatury i ewentualnie dodania elektrolitu. Fazy liotropowe tworzone są także przez wirusy mozaiki tytoniowej, makromolekuły DNA, czy nanokryształy celulozy [52].

4. Ciekłokrystaliczne kompozyty i układy dyspersyjne

Ciekle kryształy rozproszone w stałej matrycy polimerowej (PDLC, ang. *Polymer Dispersed Liquid Crystals*) są najlepiej zbadaną klasą kompozytów, które wykorzystuje się w różnego rodzaju wyświetlaczach oraz urządzeniach elektrooptycznych [55-58]. Kompozyty różnią się od próbki makroskopowej CK, ponieważ występuje w nich bardzo duża powierzchnia w stosunku do objętości, a jego własności elektrooptyczne są głównie zdeterminowane przez oddziaływanie CK z polimerem. Ponieważ jest to układ niejednorodny duże znaczenie odgrywają również oddziaływania na granicy faz. Ze względu na złożoną strukturę membrany PDLC występuje znaczna różnica między własnościami CK rozproszonego w matrycy i w próbce objętościowej. Badania NMR (NMR, ang. *Nuclear Magnetic Resonance*) [59] pozwoliły na określenie konfiguracji direktora, dynamiki molekularnej oraz zmiany rodzaju przejścia fazowego w PDLC dla krytycznej średnicy kropli nematyka równej około $0,13 \mu\text{m}$ [59, 60]. Stwierdzono także silne oddziaływanie nematyka z powierzchnią polimeru [61]. Badania dielektryczne [62] pokazały, że w kropelkach NCK 6CB rozproszonych w polimerze część molekuł zachowuje się jak w próbce objętościowej (około 10%). Molekuły te znajdują się w centrum kropelek i nie są poddane oddziaływaniu sił powierzchniowych.

Ciekle kryształy stabilizowane polimerem (PSLC, ang. *Polymer Stabilized Liquid Crystals*) lub inaczej CK modyfikowane polimerem (PMLC, ang. *Polymer Modified Liquid Crystals*) [63-65] to grupa kompozytów składających się z CK, w którym rozproszone niewielką ilości polimeru (o koncentracji mniejszej niż 10%). Po dodaniu mezogenicznego monomeru do CK powstała mieszanina znajduje się w mezofazie.

Następnie monomer poddaje się polimeryzacji w otoczeniu CK, co prowadzi do powstania sieci polimerowej, która przejmuje jego uporządkowanie i orientację. Pożądaną strukturę sieci polimerowej uzyskuje się przez nałożenie warstw orientujących i zewnętrznego pola elektrycznego, które porządkują CK w czasie procesu polimeryzacji. W ten sposób otrzymuje się określoną konfigurację direktora CK. Dzięki wykorzystaniu powyższej techniki zaobserwowano szereg bardzo interesujących efektów elektrooptycznych w stabilizowanych polimerem nematycznych [63, 66, 67], cholesterycznych [68] i ferroelektrycznych CK [69].

5. Ciekłe kryształy ograniczone geometrycznie

Materiały porowate w ostatnich latach przyciągają uwagę badaczy ze względu na ich znaczenie poznawcze i zastosowanie technologiczne. Szczególnie interesujące jest pytanie dotyczące wpływu geometrii i struktury substancji porowatej na dynamikę ruchów molekularnych i kolektywnych, własności optyczne, fluktuację parametru porządku, temperaturę przejścia fazowego, itp. cieczy umieszczonej w tego typu ośrodkach [70, 71].

Badane są materiały porowate posiadające sieć połączonych kanałów i wnęk, np. szkła Vycor [71], CPG (ang. *Controlled-Pore Glass*) [72], czy aerożel krzemionkowy (ang. *Aerogel*) posiadający trójwymiarową sieć połączonych kanałów [73]. Inne interesujące materiały porowate to membrany Anopore, Nuclepore z dobrze zdefiniowanymi cylindrycznymi porami ustawionymi prostopadle do powierzchni membrany [74], membrany Synpor z fraktalną strukturą porów [75], czy sita molekularne MCM-41 i MCM-48 [76]. Powierzchnia porów jest bardzo duża w porównaniu z objętością materiału porowatego, co czyni go bardzo użytecznym w badaniach oddziaływań z powierzchnią.

Ciekłe kryształy są dobrym kandydatem do badania wpływu ograniczeń geometrycznych na charakter oddziaływań CK-ciało stałe. Natomiast nematyki umieszczone w materiałach porowatych reprezentują prosty układ, gdzie uporządkowanie orientacyjne i oddziaływanie z powierzchnią mogą być dość łatwo zaobserwowane. Umieszczenie CK w materiale porowatym umożliwia wyjaśnienie mechanizmu kotwiczenia i porządkowania jego molekuł na powierzchni ciała stałego. Oddziaływanie to zależy zarówno od rodzaju powierzchni, jak i natury substancji ciekłokrystalicznej [16].

Nowym obszarem badań, z wieloma otwartymi pytaniami, są badania wpływu geometrii materiału porowatego na dynamikę modów kolektywnych i formowanie faz smektycznych w FCK [77-79].

5.1. Teoretyczny opis struktur porowatych

Podstawowym problemem w opisie złożonej geometrii struktur porowatych jest określenie, który z parametrów takiego układu ma znaczący wpływ na własności termodynamiczne, elektryczne, dielektryczne i optyczne umieszczonej w nim substancji. Stosowane obecnie opisy teoretyczne można podzielić na dwie kategorie: pierwsza z nich korzysta z uproszczonych modeli geometrycznych struktury porów (teoria perkolacji [80] i model ziaren [81]), natomiast druga koncentruje się w dużej mierze na założeniach fenomenologicznych [82].

W opisie klasycznym złożoną strukturę układu porowatego przybliża się przez regularne formy geometryczne. Zwykle wychodzi się od lokalnego obrazu porów,

dokonyjąc następnie jego przestrzennej transformacji. Ważnymi parametrami, które należy wziąć pod uwagę w ramach opisu klasycznego, są: rozkład rozmiarów porów, średni rozmiar porów, ścieżki perkolacyjne.

Wzrost zainteresowania substancjami porowatymi związany jest z ideą, że można je opisać jako układy zawierające samopodobne, fraktalne struktury porów [83]. Idea fizyczna jest taka, iż morfologia nieregularności powierzchniowych jest niezależna od skali powiększenia, w jakiej jest obserwowana. Zwiększając powiększenie, obserwujemy wzrost ilości nieregularności, które morfologicznie są podobne do tych obserwowanych w makroskali. Dwa parametry opisują procesy zachodzące w strukturze i dynamice fraktali: wymiar fraktalny, który związany jest z gęstością obsadzeń w strukturze fraktalnej, oraz wymiar spektralny, który pojawia się w związku z własnościami dynamicznymi fraktali i związany jest z powiązaniem dynamicznymi systemu.

5.2. Ciekłe kryształy w matrycach porowatych

Określenie warunków kotwiczenia, uporządkowania i oddziaływania molekuł CK z powierzchnią ciała stałego odgrywa podstawową rolę w wykorzystaniu tych substancji przy konstrukcji urządzeń zobrazowania informacji (displejów i macierzy ciekłokrystalicznych). Umieszczenie CK w nieorganicznej matrycy porowatej daje możliwość skonstruowania wyświetlaczy o znaczeniu porównywalnym do displejów typu PDLC [84]. Podstawowym problemem jest znalezienie materiału porowatego zapewniającego odpowiednie warunki kotwiczenia, pozwalające na przełączanie pola direktora i kontrolowanie własności optycznych układu.

Istnieje szereg technik przygotowania powierzchni ciała stałego w celu uzyskania określonego uporządkowania molekuł, np. chemiczne wytrawianie powierzchni, pokrycie jej substancją powierzchniowo czynną, jednokierunkowe zarysowanie powierzchni, czy metoda skośnego naporowania [85]. Odległość γ , na jakiej obserwuje się oddziaływanie CK z powierzchnią, jest rzędu 10^3 \AA [86]. Efekty wywołane obecnością powierzchni można także badać, umieszczając CK w materiale porowatym. Zachowanie się CK w porach o średnicy mniejszej niż γ daje szereg informacji o naturze oddziaływań CK-powierzchnia ciała stałego.

5.2.1. Ciekłe kryształy w membranach porowatych

Ciekłe kryształy w ograniczeniach geometrycznych są intensywnie badane za pomocą wielu metod, takich jak NMR [87], kalorymetria wysokiej rozdzielczości (ACC, ang. *Alternating Current Calorimetry*) [88], rezonansowe rozpraszanie promieni X (RXS, ang. *Resonant X-ray Scattering*) [89], elektronowy rezonans paramagnetyczny (EPR, ang. *Electron Paramagnetic Resonance*) [90], dynamiczne rozpraszanie światła (DLS, ang. *Dynamic Light Scattering*) [91], niskokątowe rozpraszanie neutronów [92], metody optyczne [93], szerokopasmowa spektroskopia dielektryczna (BDS, ang. *Broadband Dielectric Spectroscopy*) [94, 95].

Duża liczba prac doświadczalnych i teoretycznych została poświęcona wpływowi ograniczeń geometrycznych o symetrii cylindrycznej na różne fazy ciekłokrystaliczne. Dobrze zdefiniowana geometria porów o kształcie cylindrycznym w membranach Anopore i Nuclepore [74] stała się bodźcem do wielu badań polegających na określeniu możliwych konfiguracji direktora CK w porach, wpływowi na przejścia fazowe czy zjawisko zwilżania [96]. Z badań BDS [15] oraz NMR [61, 97] wiadomo, że direktor

NCK 5CB w niezmodyfikowanych porach wymienionych powyżej membran [97] uporządkowany jest równolegle do ich osi (PA, ang. *Parallel Axial*); po użyciu lecytyny lub kwasów alifatycznych ($C_nH_{2n+1}COOH$) jako surfaktantów do modyfikacji powierzchni, orientacja zmienia się na radialną na powierzchni, prowadząc do trzech podstawowych typów konfiguracji direktora nematycznego: planarno-radialnej (PR, ang. *Planar Radial*), planarno-polarnej (PP, ang. *Planar Polar*) i radialnej z deformacją w kierunku osi cylindra (ER, ang. *Escaped Radial*). Orientacja homeotropowa (molekuły ustawione prostopadłe do ścianek porów) była stabilna dla $n > 7$, natomiast dla łańcuchów o długości $n < 6$ otrzymywano orientację planarną (molekuły ustawione równolegle do ścianek porów). Oddziaływanie między długimi łańcuchami alkilowymi a molekułami CK ma naturę steryczną i prowadzi do orientacji homeotropowej, natomiast dla krótkich łańcuchów dominuje silne oddziaływanie między powierzchnią matrycy porowatej i rdzeniem CK, co prowadzi do orientacji planarnej. Analiza teoretyczna wykazała, że struktura PP jest preferowana w porach o dużym stopniu zakrzywienia lub wystarczająco słabych warunkach kotwiczenia na ściankach wnęki. Jednakże, dla wystarczająco dużych promieni lub silnych warunków kotwiczenia faworyzowana jest struktura typu ER [97].

Badania ciepła właściwego w przejściu fazowym dla CK z grupy n-cyjanobifenyli (nCB, n – liczba atomów węgla) umieszczonych w membranach Anopore pokazały, że przejście fazowe nematyk-faza izotropowa (NI) jest silnie stłumione, poszerzone i przesunięte w stronę niższych temperatur [98, 99]. Co więcej, przejście fazowe smektyk A-nematyk (SmA-N) dla 8CB w Anopore jest znacząco bardziej zaburzone niż przejście fazowe NI. Natomiast przejście SmA-I jest bardzo poszerzone i większość efektów obserwowanych dla przejść fazowych NI i SmA-N nie występuje. Ponadto dzięki pomiarom ACC oraz RXS zostało stwierdzone dyskretne, smektyczne uporządkowanie w fazie izotropowej dla 12CB w membranach Anopore [100]. Zwiększanie się uporządkowania smektycznego jest wzmacniane przez powierzchnię porów po jej pokryciu kwasami alifatycznymi o zmiennej długości łańcucha. Nie zauważono tego efektu w porach o niezmodyfikowanej powierzchni lub pokrytych lecytyną, jak również dla CK, które posiadają fazę nematyczną. Jednocześnie w innych nCB odkryto nowe efekty związane z modyfikacją wielkości ciepła właściwego, przesunięciem temperatury przejścia fazowego oraz *wyglądaniem* przejścia fazowego [98].

Obserwacje transmisji światła spolaryzowanego przechodzącego przez membrany wypełnione NCK pokazują, że w wyniku występowania dwójłomności materiału ciekłokrystalicznego membrana Anopore jest widoczna lub nie w zależności od konfiguracji molekuł w porach. W nieobecności surfaktantu widoczny jest czarny krzyż Maltański charakterystyczny dla kryształów jednoosiowych wyciętych prostopadłe do osi optycznej. W membranach o powierzchni zmodyfikowanej przez surfaktant (lecytyna, kwasy alifatyczne) indukowana jest homeotropowa orientacja direktora na ściankach i obserwuje się silne rozpraszanie światła. Efektywny współczynnik załamania, który nie jest dopasowany do współczynnika załamania membrany, powoduje, że staje się ona mętną.

Badanie dynamiki molekularnej NCK 5CB umieszczonego w membranie Anopore o średnicy porów rzędu 200 nm przeprowadzono z wykorzystaniem BDS [15, 75, 77, 78, 101]. W zależności od przygotowania powierzchni porów, molekuły CK ustawiają się osiowo (równolegle do osi porów) lub radialnie (prostopadłe do osi porów). Orientacja planarna (osiowa) była zawsze obserwowana w membranach bez obróbki powierzchni

porów. Natomiast orientację homeotropową (radialną) uzyskiwano, pokrywając powierzchnię porów kwasem dekanokarboksylowym $C_{10}H_{19}COOH$. Zaobserwowano, że dynamika relaksacji procesu związanego z rotacją molekuly wokół krótkiej osi (proces δ) oraz procesu libracyjnego związanego z oscylacją wokół kierunku direktora ulega modyfikacji. W stosunku do próbki objętościowej częstotliwość charakterystyczna procesu δ rośnie, a procesu libracyjnego maleje. Zaobserwowane prawidłowości powiązane ze zmianą parametru porządku. Ponadto wyznaczono anizotropię dielektryczną dla 5CB oraz zaobserwowano podwójną zmianę znaku przenikalności elektrycznej [15]. Podobne prawidłowości zaobserwowano w układzie membrana Synpor-NCK 5CB [75].

5.2.2. Ciekłe kryształy ze strukturą helikoidalną w matrycach porowatych

W przypadku chiralnych faz ciekłokrystalicznych umieszczonych w porach, orientacja planarna lub homeotropowa na ściankach konkuruje z niezaburzoną strukturą helikoidalną wewnątrz porów. W tej sytuacji obserwuje się znacznie większą liczbę defektów powierzchniowych i objętościowych w badanym układzie.

Własności optyczne chiralnych CK umieszczonych w membranach Anopore ulegają zmianie pod wpływem geometrii matrycy porowatej. Znak aktywności optycznej dla smektyka C^* umieszczonego w membranie Anopore jest odwrócony w porównaniu do próbki objętościowej. Dla układu cholesteryk-matryca porowata Anopore stwierdzono rotację płaszczyzny polaryzacji w prawo dla wszystkich badanych temperatur. W przypadku Ch i SmC^* w membranie Anopore zaobserwowano także dodatnią dwójłomność, która rosła wraz ze spadkiem temperatury [102]. Doświadczalne badania chiralnych nematyków (np. 5CB z domieszką chiralną CB15 i CE2) umieszczonych w submikronowych cylindrycznych wnękach pokazują, że następuje przejście konfiguracyjne z uporządkowania osiowego do planarnego w wyniku zmniejszania długości skoku helisy mieszaniny chiralnego nematyka [103].

Badania dwójłomności optycznej, dyfrakcji neutronów oraz własności elektrooptycznych FCK 2MBOCBC (ang. *S-(-)-2-methylbutyl 4-n-nonanoyloxy-biphenyl-4'-carboxylate*) umieszczonego w porowatych membranach AAO (ang. *Anodic Aluminium Oxide Monoliths*) z anodowanego tlenku glinu o średnicy porów 10, 15 i 20 nm pokazały, że umieszczenie FCK w porach bez modyfikacji powierzchni nie stanowi dobrego ośrodka dla stabilnego kotwiczenia molekuly w fazie smektycznej. Dopiero pokrycie ścianek porów powłoką polimerową zwiększa planarne kotwiczenie molekuly na ich powierzchni. Powoduje to formowanie fazy SmC^* o dużej aktywności optycznej, której oś helisy jest równoległa do osi porów. W badanym układzie zaobserwowano GM, którego częstotliwość rosła wraz z kwadratem odwrotności promienia porów [104].

Z badań BDS FCK 2MBOCBC umieszczonego w membranach Anopore wynika, że ze względu na idealne uporządkowanie warstw smektycznych w porach o symetrii cylindrycznej (dla FCK o dużym skoku helikoidy i małej wartości polaryzacji spontanicznej) mody kolektywne są nieaktywne, ponieważ występuje brak sprzężenia między polaryzacją spontaniczną warstw smektycznych a mierzącym polem elektrycznym [105]. Co ciekawe, wyniki pomiarów temperaturowej zmiany inkrementu dielektrycznego w przejściu fazowym z fazy SmA do fazy SmC^* wykazały, że faza smektyczna jest nadal obecna w porach membrany Anopore, mimo że nie obserwuje się GM [106]. Natomiast zanik tego modu w membranach Synpor o złożonej strukturze porów wynika z wpływu zamrożonego przypadkowego nieporządku defektów na formowanie struktury helikoidalnej w fazie SmC^* [105].

6. Wnioski

Przedstawiony przegląd wyników badań pokazuje, że domieszkowanie ciekłych kryształów nanocząstkami różnych materiałów prowadzi do modyfikacji ich własności elektrooptycznych, dielektrycznych, przewodnictwa, lepkości, uporządkowania molekuł oraz wielu innych interesujących i często niespodziewanych zjawisk, które są pożądane z punktu widzenia zastosowań technologicznych. Istotnym czynnikiem w uzyskaniu nowych materiałów nanostrukturalnych jest zastosowanie nanocząstek o określonej geometrii, rozmiarach i własnościach powierzchni. Należy podkreślić, że jednym z celów najnowszych badań ciekłych kryształów jest dążenie do rozszerzenia ich własności metodami fizycznymi bez potrzeby syntezy chemicznej nowych materiałów.

Do modyfikacji własności fizycznych ciekłych kryształów wykorzystuje się między innymi nanocząstki ferroelektryczne, nanocząstki metali szlachetnych, nanocząstki półprzewodników, czy nanocząstki węgla. Okazuje się, że wprowadzenie nanocząstek magnetycznych do ciekłych kryształów prowadzi do powstania materiałów ferronematycznych o zaskakujących własnościach fizycznych. Domieszkowanie ciekłych kryształów nanocząstkami ferroelektrycznymi umożliwia uzyskanie efektu pamięci. Umieszczenie nanokropek kwantowych w cholesterycznych ciekłych kryształach pozwala na budowę przestrajalnych laserów nowej generacji. Anizotropowe własności ciekłych kryształów umożliwiają przestrzenne porządkowanie kropek kwantowych oraz innych nanocząstek, prowadząc do uzyskania metamateriałów o niespotykanych wcześniej własnościach optycznych. Lokowanie się nanocząstek w liniach dysklinacji umożliwia stabilizację i rozszerzenie zakresu temperaturowego fazy niebieskiej w cholesterycznych ciekłych kryształach, co prowadzi do potencjalnego zastosowania tych materiałów w konstrukcji displejów nowej generacji. Ponadto efekt samoorganizacji nanocząstek w liotropowe fazy ciekłokrystaliczne o intrygujących, nowych własnościach może znaleźć zastosowanie w nowej klasie przyrządów fotonicznych. Natomiast obecność nanocząstek aerosilu w strukturze nematycznego oraz ferroelektrycznego ciekłego kryształu prowadzi do modyfikacji własności elektrooptycznych (efekt pamięci) oraz dynamiki procesów molekularnych i kolektywnych. W ciekłych kryształach ograniczonych geometrycznie własności dielektryczne oraz dynamika procesów molekularnych i kolektywnych ulegają również znacznej modyfikacji.

Bardzo dynamicznie rozwijającą się dziedziną, której nie uwzględniono w tym przeglądzie, jest wykorzystanie odpowiednio sfunckjonalizowanych nanocząstek do przenoszenia leków oraz biosensorów ciekłokrystalicznych zmodyfikowanych nanocząstkami do badania procesów zachodzących w organizmie na poziomie molekularnym.

Literatura

1. An J.-G., Hina S., Yong Y., Xue M., Liu Y., *Characterization of liquid crystals: a literature review*, Review of Advanced in Materials Science, 44, 2016, s. 398-406.
2. Chen H.-W., Lee J.-H., Lin B.-Y., Chen S., Wu S.-T., *Liquid crystal display and organic light-emitting diode display: Present status and future perspectives*, Light: Science & Applications, 7, 2018, s. 17168.
3. Smith C.A., *A review of liquid crystal display technologies, electronic interconnection and failure analysis*, Circuit World, 34, 2008, s. 35-41.
4. Miniewicz A., Gniewek A., Parka J., *Liquid crystals for photonic applications*, Optical Materials, 21, 2002, s. 605-610.

5. Lagerwall J.P.F., Scalia G., *A new era for liquid crystal research: Applications of liquid crystals in soft matter nano-, bio- and microtechnology*, Current Applied Physics, 12, 2012, s. 1387-1412.
6. Wongkaew N., Simsek M., Criesche Ch., Baeumner A.J., *Functional nanomaterials and nanostructures enhancing electrochemical biosensors and lab-on-a-chip performances: Recent progress, applications, and future perspective*, Chemical Review, 119, 2019, s. 120-194.
7. Bisoyia H.K., Kumar S., *Liquid-crystal nanoscience: An emerging avenue of soft self-assembly*, Chemical Society Review, 40, 2011, s. 306-319.
8. Manjuladevi V., Guptas R., *Effect of nanomaterials on physical properties of liquid crystals*, [w:] *New development in liquid crystals research*, Nova Publishers, USA 2015.
9. Yadav S., Malik P., *Effect of size and concentration of magnetic nanoparticles on blue phase stabilization and electro-optical properties in blue phase liquid crystalline nanocomposites*, Optical Materials, 122, 2021, s. 111670.
10. Róžański S.A., Sinha G.P., Thoen J., *Influence of hydrophilic and hydrophobic aerosil particles on the molecular modes in the liquid crystal 4-n-pentyl-4'-cyanobiphenyl*, Liquid Crystals, 33, 2006, s. 833-840.
11. Róžański S.A., Thoen J., *Collective dynamic modes in ferroelectric liquid crystal-aerosil dispersions*, Liquid Crystals, 32, 2005, s. 331-341.
12. Li Q. (red.), *Nanoscience with liquid crystals*, Springer Cham, Switzerland 2014.
13. Stamatoiu O., Mirzaei J., Feng X., Hegmann T., *Nanoparticles in liquid crystals and liquid crystalline nanoparticles*, Liquid Crystals, 318, 2011, s. 331-393.
14. Polarza S., Smarsly B., *Nanoporous materials*, Journal of Nanoscience and Nanotechnology, 2, 2002, s. 581-612.
15. Róžański S.A., Stannarius R., Groothues H., Kremer F., *Dielectric properties of the nematic liquid crystal 4-n-pentyl-4'-cyanobiphenyl in porous membranes*, Liquid Crystals, 20, 1996, s. 59-66.
16. Crawford G.P., Žumer S. (red.), *Liquid crystals in complex geometries*, Taylor and Francis, London 1996.
17. Singh S., *Liquid crystals: Fundamentals*, World Scientific, 2002.
18. Chandrasekhar S., Hull M.D., *Liquid crystals*, Cambridge University Press, 1992.
19. Oswald P., Pierański P., *Smectic and columnar liquid crystals*, CRC Press, 2005.
20. Atkins P.W., Friedman R.S., *Molecular quantum mechanics*, OUP, Oxford 2011.
21. Lagerwall S.T., *Ferroelectric and antiferroelectric liquid crystals*, Wiley & Sons, 2008.
22. Stegemeyer H. (red.), *Liquid crystals*, Springer Science & Business Media, 2013.
23. Figueiredo Neto A.M., Salinas S.R.A., *The physics of lyotropic liquid crystals: Phase transitions and structural properties*, Oxford University Press, 2005.
24. Bleeker E.A.J., de Jong W.H., Geertsma R.E., Groenewold M., Heugens E.H.W., Koers-Jacquemijns M., de Meent D., Popma J.R., Rietveld A.G., Wijnhoven S.W.P., Cassee F.R., Oomen A.G., *Considerations on the EU definition of a nanomaterial: Science to support policy making, regulatory toxicology and pharmacology*, Regulatory Toxicology and Pharmacology, 65, 2013, s. 119-125.
25. Świdarska-Sroda A., Łojkowski W., Lewandowska M., Kurzydłowski K.J. (red.), *Świat nanoczqstek*, PWN, Warszawa 2016.
26. Hegmann T., Qi H., Marx V.M., *Nanoparticles in liquid crystals: Synthesis, self-assembly, defect formation and potential applications*, Journal of Inorganic and Organometallic Polymers and Materials, 17, 2007, s. 483-508.
27. Shen Y., Dierking I., *Perspectives in liquid-crystal-aided nanotechnology and nanoscience*, Applied Sciences, 9, 2019, s. 2512-2559.
28. <https://www.szfki.hu/~compli/Ferronematics/Ferronematics.htm> [data dostępu: 20.09.2022].

29. Brochard F., De Gennes P., *Theory of magnetic suspensions in liquid crystals*, Journal de Physique, 31, 1970, s. 691-708.
30. Kopčanský P., Tomašovičová N., Koneracká M., Závišová V., Timko M., Džarová A., Šprincová A., Éber N., Fodor-Csorba K., Tóth-Katona T., Vajda A., Jadżyn J., *Structural changes in the 6CHBT liquid crystal doped with spherical, rodlike, and chainlike magnetic particles*, Physical Review E, 78, 2008, 011702/1-5.
31. Csach K., Juríková A., Miškuf J., Tomašovičová N., Gdovinová V., Závišová V., Kopčanský P., Éber N., Fodor-Csorba K., Vajda A., *Kinetics of nematic to isotropic phase transition in liquid crystal doped with magnetic nanoparticles*, Acta Physica Polonica, 131, 2017, s. 949-951.
32. Mitrooova Z., Tomašovičová N., Koneracká M., Závišová V., Timko M., Kovač J., Kopčanský P., Tomčo L., Éber N., Fodor-Csorba K., Tóth Katona T., Vajda A., Jadżyn J., Beaugnon E., Chaud X., *The structural transitions in a nematics liquid crystal doped with magnetically labeled carbon nanotubes*, Magnetohydrodynamics, 45, 2009, s. 353-360.
33. Kempaiah R., Liu Y., Nie Z., Basu R., *Giant soft-memory in liquid crystal nonocomposites*, Applied Physics Letters, 108, 2016, s. 83105.
34. Basu R., *Soft memory in a ferroelectric nanoparticle-doped liquid crystal*, Physical Review E, 89, 2014, s. 22508.
35. Glushchenko A., Kresse H., Puchkovska G., Reshetnyak V., Reznikov Yu., Yaroshchuk O., *Memory Effect and structure of filled nematic liquid crystals*, Molecular Crystals and Liquid Crystals, 321, 1998, s. 15-30.
36. Eidenschink R., de Jeu W.H., *Static scattering in filled nematic: New liquid crystal display technique*, Electronics Letters, 27, 1991, s. 1195-1196.
37. Kreuzer M., Tschudi T., Eidenschink R., *Erasable optical data storage in bistable liquid crystal cells*, Molecular Crystals and Liquid Crystals, 223, 1992, s. 219-227.
38. Gdovinová V., Tomašovičová N., Shie-Chan Jeng, Zakutanská K., Kula P., Kopčanský P., *Memory effect in nematic phase of liquid crystal doped with magnetic and non-magnetic nanoparticles*, Journal of Molecular Liquids, 282, 2019, s. 286-291.
39. Kumar A., Priyam Meena H., Prakash J., Wang L., Singh G., *Recent advances on semiconducting nanomaterials – ferroelectric liquid crystals nanocomposites*, Journal of Physics: Condensed Matter, 34, 2022, s. 13004.
40. Cho Ch.-Y., Lee S.-J., Song J.-H., Hong S.-H., Lee S.-M., Cho Y.-H., Park S.-J., *Enhanced optical output power of green light-emitting diodes by surface plasmon of gold nanoparticles*, Applied Physics Letters, 98, 2011, s. 51106.
41. Rodarte A.L., Pandolfi R.J., Ghosh S., Hirst L.S., *Quantum dot/liquid crystal composite materials: self-assembly driven by liquid crystal phase transition templating*, Journal of Materials Chemistry C, 1, 2013, s. 5527-5532.
42. Chen L.-J., Lin J.-D., Lee C.-R., *An optically stable and tunable quantum dot nanocrystal-embedded cholesteric liquid crystal composite laser*, Journal of Materials Chemistry C, 2, 2014, s. 4388-4394.
43. Choudharya A., Singha A.G., Biradar A.M., *Advances in gold nanoparticle-liquid crystal composites*, Nanoscale, 14, 2014, s. 7743-7756.
44. Podgornov F.V., Ryzhkova A.V., Haase W., *Influence of gold nanorods size on electro-optical and dielectric properties of ferroelectric liquid crystals*, Applied Physics Letters, 97, 2010, s. 212903.
45. Yoshida H., Tanaka Y., Kawamoto K., Kubo H., Tsuda T., Fujii A., Kuwabata S., Kikuchi H., Ozaki M., *Nanoparticle-stabilized cholesteric blue phases*, Applied Physics Express, 2, 2009, s. 121501.
46. Kikuchi H., Yokota M., Hisakado Y., Yang H., Kajiyama T., *Polymer-stabilized liquid crystal blue phases*, Nature Materials, 1, 2002, s. 64-68.

47. Wong J.-M., Hwanga J.-Y., Chien L.-C., *Electrically reconfigurable and thermally sensitive optical properties of gold nanorods dispersed liquid crystal blue phase*, *Soft Matter*, 18, 2011, s. 7956-7959.
48. Lisetski L., Soskin M., Lebovka N., *Carbon nanotubes in liquid crystals: Fundamental properties and applications*, [w:] *Physics of liquid matter: Modern problems*, Springer Proceedings in Physics, V. 171, Springer International Publishing, Switzerland 2015, s. 243-297.
49. Lagerwall J.P.F., Scalia G., *Carbon nanotubes in liquid crystals*, *Journal of Materials Chemistry*, 18, 2008, s. 2890-2898.
50. Zhang X., Yu B., Cong H., *Recent developments in fullerene-containing thermotropic liquid crystals*, *Current Organic Chemistry*, 21, 2017, s. 1600-1611.
51. Lekkerkerker H.N.W., Vroege G.J., *Liquid crystal phase transitions in suspensions of mineral colloids: New life from old roots*, *Philosophical Transactions of the Royal Society A*, 371, 2013, 20120263.
52. Dierking I., Figueiredo Neto A.M., *Novel trends in lyotropic liquid crystals*, *Crystals*, 10, 2020, s. 604-628.
53. Onsager L., *The effects of shape on the interaction of colloidal particles*, *Annals of the New York Academy of Sciences*, 51, 1949, s. 627-659.
54. Puech N., Blanc C., Grelet E., Zamora-Ledezma C., Maugey M., Zakri C., Anglaret E., Poulin P., *Highly ordered carbon nanotube nematic liquid crystals*, *Journal of Physical Chemistry C*, 115, 2011, s. 3272-3278.
55. Saeed M.H., Zhang S., Cao Y., Zhou L., Hu J., Muhammad I., Xiao J., Zhang L., Yang H., *Recent advances in the polymer dispersed liquid crystal composite and its applications*, *Molecules*, 25, 2020, s. 5510.
56. Bronnikov S., Kostromin S., Zuev V., *Polymer-dispersed liquid crystals: Progress in preparation, investigation, and application*, *Journal of Macromolecular Science, Part B*, 52, 2013, s. 1718-1735.
57. Drzaic P., Drzaic P.S., *Putting liquid crystal droplets to work: A short history of polymer dispersed liquid crystals*, *Liquid Crystals*, 33, 2006, s. 1281-1296.
58. Hakemi H., Pinshow O., Gal-Fuss D., *Evaluation of polymer dispersed liquid crystal (PDLC) for passive rear projection screen application*, *Recent Progress in Materials*, 1, 2019, s. 1-15.
59. Golemme A., Žumer S., Allender D.W., Doane J.W., *Continuous nematic-isotropic transition in submicron-size liquid-crystal droplets*, *Physical Review Letters*, 61, 1988, s. 2937-2940.
60. Vilfan I., Vilfan M., Žumer S., *Orientalional order in bipolar nematic microdroplets close to the phase transition*, *Physical Review A*, 40, 1989, s. 4724-4730.
61. Crawford G.P., Stannarius R., Doane J.W., *Surface-induced orientational order in the isotropic phase of a liquid-crystal material*, *Physical Review A*, 44, 1991, s. 2558-2569.
62. Jadzyn J., *Dielectric relaxation in polymer dispersed nematic liquid crystal films*, *Liquid Crystals*, 26, 1999, s. 453-456.
63. Broer D., Gossink R., Hikmet R., *Oriented polymer networks obtained by photopolymerization of liquid-crystalline monomers*, *Die Angewandte Makromolekulare Chemie*, 183, 1990, s. 45-65.
64. Hikmet R.A.M., *Anisotropic gels and plasticized networks formed by liquid crystal molecules*, *Liquid Crystals*, 9, 1991, s. 405-416.
65. Derking I. (red.), *Polymer-modified liquid crystals*, Royal Society of Chemistry, 2019.
66. Hikmet R.A.M., *Electrically induced light scattering from anisotropic gels*, *Journal of Applied Physics*, 68, 1990, s. 4406-4412.
67. Hikmet R.A.M., De Witz C., *Gel layers for inducing adjustable pretilt angles in liquid system*, *Journal of Applied Physics*, 70, 1991, s. 1265-1269.

68. Zhao D.-X., Jiang Q., Wang J., Qiu Y., Liao Y.-G., Xie X.-L., *Visible light and temperature regulated reflection colors in self-supporting cholesteric liquid crystal physical gels*, Chinese Journal of Polymer Science, 39, 2021, s. 1617-1625.
69. Derking I., *A review of polymer-stabilized ferroelectric liquid crystals*, Materials (Basel), 7, 2014, s. 3568-3587.
70. Awschalom D.D., Drake J.M., Klafter K., Kopelman R. (red.), *Dynamics in Small Confining Systems*, MRS Symposium Proceedings, V. 290, Materials Research Society, Boston 1993.
71. McKenna G.B., *Ten (or more) years of dynamics in confinement: Perspectives for 2010*, European Physical Journal Special Topics, 189, 2010, s. 285-302.
72. Gudimall A., Thomas S., Zidanšek A., *Phase behaviour of n-CB liquid crystals confined to controlled pore glasses*, Journal of Molecular Structure, 1235, 2021, s. 130217.
73. Bellini T., Clark N.A., Muzny Ch.D., Wu L., Garland C.W., Schaefer D.W., Oliver B.J., *Phase behavior of the liquid crystal 8CB in a silica aerogel*, Physical Review Letters, 69, 1992, s. 788-791.
74. Crawford G.P., Steele L.M., Ondris-Crawford R., Iannacchione G.S., Yeager C.J., Doane J.W., Finotello D., *Characterization of the cylindrical cavities of Anopore and Nuclepore membranes*, Journal of Chemical Physics, 96, 1992, s. 7788-7796.
75. Róžański S.A., *The molecular dynamics of nematic liquid crystal confined in porous membranes treated by different surfactants*, Opto-Electronics Review, 27, 2019, s. 339-344.
76. Frunza S., Frunza L., Schönhals A., Zubowa H.-L., Kosslick H., Cariusc H.-E., Fricke R., *On the confinement of liquid crystals in molecular sieves: dielectric measurements*, Chemical Physics Letters, 307, 1999, s. 167-176.
77. Kremer F., Huwe A., Schönhals A., Róžański S.A., *Molecular dynamics in confining space*, [w:] Kremer F., Schönhals A. (red.), *Broadband dielectric spectroscopy*, Springer-Verlag, New York 2002, s. 171-224.
78. Róžański S.A., *Dynamics of molecular and collective relaxation processes in confined liquid crystals*, [w:] Galewski Z., Sobczyk L. (red.), *Dielectric properties of liquid crystals*, Research Signpost, 2008, s. 183-216.
79. Kuczyński W., Hoffman J., Małecki J., *Investigation of ferroelectric modes in liquid crystals using dielectric and optical methods*, Ferroelectrics, 150, 1993, s. 279-290.
80. Hunt A.G., Sahimi M., *Flow, transport, and reaction in porous media: Percolation scaling, critical-path analysis, and effective medium approximation*, Reviews of Geophysics, 55, 2017, s. 993-1078.
81. Mecke K., Arns C.H., *Fluids in porous media: A morphometric approach*, Journal of Physics: Condensed Matter, 17, 2005, s. 503-534.
82. Gosh K., Fuchs R., *Spectral theory for two-component porous media*, Physical Review B, 38, 1988, s. 5222-5236.
83. Ge X., Fan Y., Deng S., Han T., Liu J., *An improvement of the fractal theory and its application in pore structure evaluation and permeability estimation*, Journal of Geophysical Research: Solid Earth, 121, 2016, s. 6333-6345.
84. He Z., Yin K., Wu S.-T., *Passive polymer-dispersed liquid crystal enabled multi-focal plane displays*, Optics Express, 28, 2020, s. 15294-15299.
85. Sheng P., *Phase transition in surface-aligned nematic films*, Physical Review Letters, 37, 1976, s. 1059-1062.
86. Blinov L.M., Kac E.I., Sonin A.A., *Surface physics of thermotropic liquid crystals*, Soviet Physics Uspekhi, 30, 1987, s. 604-619.
87. Jin T., Finotello D., *Controlling disorder in liquid crystal aerosol dispersions*, Physical Review E, 69, 2004, s. 41704.
88. Roshi A., Iannacchione G.S., Clegg P.S., Birgeneau R.J., *Evolution of the isotropic-to-nematic phase transition in octyloxycyanobiphenyl+aerosol dispersions*, Physical Review E, 69, 2004, s. 31703.

89. Cordoyiannis G., Nounesis G., Bobnar V., Kralj S., Kutnjak Z., *Confinement-induced orientational order in a ferroelectric liquid crystal containing dispersed aerosils*, Physical Review Letters, 94, 2005, s. 27801.
90. Arcioni A., Bacchiocchi I., Giampaolo V., Venditti G., Zannon C., *A comparison of the effects of dispersed hydrophobic or hydrophilic aerosol nanoparticles on the order and dynamics of the 5CB liquid crystal*, Chemical Physics Letters, 396, 2004, s. 433-441.
91. Aliev F., *Molecular and collective relaxation in deeply supercooled confined liquid crystal*, Journal of Non-Crystalline Solids, 489, 2002, s. 307-310.
92. Jáklí A., Almásy L., Borbély S., Rosta L., *Memory of silica aggregates dispersed in smectic liquid crystals: Effect of the interface properties*, European Physical Journal B, 10, 1999, s. 509-513.
93. Binder H., Schmiedel H., Lantzsch G., Cramer C., Klose G., *Molecular ordering in microconfined liquid crystals: An infrared linear dichroism study*, Liquid Crystals, 21, 1996, s. 415-426.
94. Róžański S.A., Thoen J., *Dynamics of the Goldstone mode near the point of polarization sign reversal in ferroelectric liquid crystal–aerosil mixtures*, Liquid Crystals, 32, 2005, s. 1013-1020.
95. Kremer F. (red.), *Dynamics in geometrical confinement*, Springer International Publishing, 2014.
96. Liu X., Finotello D., *Liquid crystal wetting in cylindrical pores*, Molecular Crystals and Liquid Crystals, 367, 2001, s. 341-351.
97. Crawford G.P., Allender D.W., Doane J.W., *Surface elastic and molecular-anchoring properties of nematic liquid crystals confined to cylindrical cavities*, Physical Review A, 45, 1992, s. 8693-8708.
98. Iannacchione G.S., Finotello D., *Calorimetric study of phase transitions in confined liquid crystals*, Physical Review Letters, 69, 1992, s. 2094-2097.
99. Iannacchione G.S., Finotello D., *Specific heat dependence on orientational order at cylindrically confined liquid crystal phase transitions*, Physical Review E, 50, 1994, s. 4780-4795.
100. Iannacchione G., Mang J.T., Kumar S., Finotello D., *Surface-induced discrete smectic order in the isotropic phase of 12CB in cylindrical pores*, Physical Review Letters, 73, 1994, s. 2708-2711.
101. Róžański S.A., Kremer F., Groothues H., Stannarius R., *The dielectric properties of nematic liquid crystal 5CB confined to treated and untreated Anopore membranes*, Molecular Crystals and Liquid Crystals, 303, 1997, s. 319-324.
102. Schmiedel H., Stannarius R., Feller G., Cramer Ch., *Experimental evidence of a conic helical liquid crystalline structure in cylindrical microcavities*, Molecular Crystals and Liquid Crystals, 282, 1996, s. 395-405.
103. Ondris-Crawford R.J., Ambrožič M., Doane J.W., Žumer S., *Pitch-induced transition of chiral nematic liquid crystals in submicrometer cylindrical cavities*, Physical Review E, 50, 1994, s. 4773-4779.
104. Busch M., Kityk A.V., Piecek W., Hofmann T., Wallacher D., Caľus S., Kula P., Steinhart M., Eich M., Huber P., *A ferroelectric liquid crystal confined in cylindrical nanopores: Reversible smectic layer buckling, enhanced light rotation and extremely fast electro-optically active Goldstone excitations*, Nanoscale, 9, 2017, s. 19086-19099.
105. Róžański S.A., Stannarius R., Kremer F., *Collective dynamic modes of microconfined ferroelectric liquid crystals*, IEEE Transactions on Dielectrics and Electrical Insulation, 8, 2001, s. 488-493.
106. Róžański S.A., Thoen J., *Evidence for a tilted smectic phase in Anopore membranes by dielectric spectroscopy*, Liquid Crystals, 33, 2006, s. 1043-1049.

Ciekłe kryształy domieszkowane nanocząstkami i ograniczone geometrycznie – nowe zjawiska fizyczne i możliwości technologiczne

Streszczenie

W pracy pokazano, że ciekłe kryształy w połączeniu z nanomateriałami prowadzą do powstania całkowicie nowych, innowacyjnych materiałów, które znajdują obecnie coraz większe zastosowanie w takich dziedzinach, jak optoelektronika, obrazowanie informacji, elektrooptyka, itp. Modyfikacja struktury molekularnej ciekłych kryształów poprzez dodanie nanocząstek metali, półprzewodników, węgla, itp. lub umieszczenie ich w matrycach porowatych objawia się bogactwem wielu nowych efektów i zjawisk fizycznych oraz prowadzi do modyfikacji własności optycznych, dielektrycznych, czy przewodnictwa kompozytów ciekło-krysztalicznych. Szczególną uwagę zwrócono na możliwości otrzymania nowych materiałów poprzez domieszkowaniem ciekłych kryształów nanocząstkami złota, nanorurkami węglowymi, nanocząstkami krzemionki oraz samoorganizację wydłużonych, nieorganicznych nanocząstek w struktury analogiczne do liotropowych ciekłych kryształów. Wprowadzenie nanocząstek do ciekłego kryształu przyczynia się do stabilizacji faz ciekłokrystalicznych oraz umożliwia konstrukcję przestrajalnych laserów. Dynamika relaksacji procesów molekularnych i kolektywnych w ciekłokrystalicznych układach dyspersyjnych oraz w ciekłych kryształach ograniczonych geometrycznie ulega znacznej modyfikacji w wyniku wprowadzenia do mezofazy kontrolowanego nieporządku, oddziaływania molekuł z powierzchnią oraz efektów skończonych rozmiarów. Dziedzina nanokompozytów ciekłokrystalicznych rozwija się bardzo dynamicznie, a najbliższe lata powinny zaowocować odkryciem i wdrożeniem wielu nowych technologii i urządzeń wykorzystujących własności tego typu układów dyspersyjnych.

Słowa kluczowe: ciekłe kryształy, nanocząstki, ograniczenia geometryczne, nanotechnologia, samoorganizacja

Liquid crystals doped with nanoparticles and geometrically constrained – new physical phenomena and technological possibilities

Abstract

The paper shows that liquid crystals in combination with nanomaterials lead to the creation of completely new, innovative materials that are now increasingly used in such fields as optoelectronics, information imaging, electro-optics, etc. Modification of the molecular structure of liquid crystals by adding metal nanoparticles, semiconductors, carbon, etc. or confining liquid crystals in porous matrices reveals a wealth of many new physical effects and phenomena and leads to modification of optical, dielectric or conductivity properties of liquid crystal composites. Particular attention was paid to the possibility of obtaining new materials by doping liquid crystals with gold nanoparticles, carbon nanotubes, silica nanoparticles and self-assembly of elongated, inorganic nanoparticles into structures analogous to the lyotropic liquid crystals. The introduction of nanoparticles into the liquid crystal contributes to the stabilization of the liquid crystal phases or enables the construction of tunable lasers. The relaxation dynamics of molecular and collective processes in liquid crystal dispersions and in geometrically restricted liquid crystals is significantly modified as a result of introducing a controlled disorder into the liquid crystal, the interaction of molecules with the surface and the finite size effects. The field of liquid crystal nanocomposites is developing very dynamically, and the coming years should result in the discovery and implementation of many new technologies and devices using the properties of this type of dispersion systems.

Keywords: liquid crystals, nanoparticles, geometric constraints, nanotechnology, self-organization

Charakterystyka spektroskopowa oraz właściwości aplikacyjne nowych up-konwersyjnych nanomateriałów nieorganicznych

1. Cel pracy

Nanotechnologia jako intensywnie rozwijająca się dziedzina interdyscyplinarna, o ogromnym potencjale aplikacyjnym, rozbudziła zainteresowania grup badawczych na całym świecie. Szerokie możliwości projektowania nanomateriałów poprzez nadawanie im pożądanych właściwości, np. fizykochemicznych, czy biologicznych, pozwoliły na prowadzenie badań, łącząc różne dziedziny nauki, co w efekcie wpłynęło na rozwój współczesnych technologii. Najnowsze doniesienia literaturowe wskazują na gwałtowny zainteresowanie tematyką otrzymywania nanomateriałów domieszkowanych jonami lantanowców wykazującymi zjawisko up-konwersji, stosowanych m.in. w termometrii optycznej, ogniwach słonecznych czy medycynie, które omówiono w niniejszej pracy.

2. Wprowadzenie

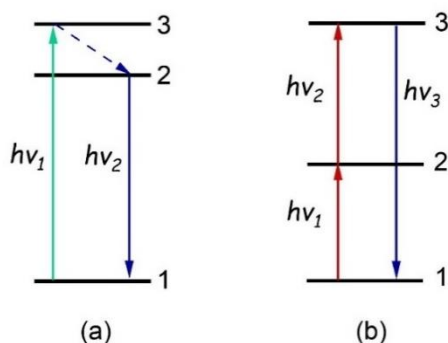
Up-konwersja (ang. *up-conversion*, UC) to nieliniowy proces optyczny polegający na konwersji energii w górę. Wzbudzając niskoenergetycznym promieniowaniem z zakresu bliskiej podczerwieni (NIR) lub podczerwieni (IR), obserwuje się emisję światła widzialnego (Vis) lub ultrafioletowego (UV), co oznacza, że następuje konwersja na promieniowanie wysokoenergetyczne. Up-konwersja nazywana jest również emisją anty-Stokesowską, ze względu na fakt, że długość fali wzbudzenia jest większa od długości fali emisji. W 1959 roku Nicolaas Bloembergen z Harvard University opublikował pierwsze doniesienia o możliwości wykrycia oraz zliczania fotonów w podczerwieni (ang. *Infrared quantum counter*, IRQC) w ciele stałym za pomocą detektora [1]. Natomiast w 1966 roku Francois Auzel zaproponował mechanizm transferu energii zachodzący pomiędzy stanami wzbudzonymi jonów w procesie up-konwersji na przykładzie luminoforu CaWO_4 współdomieszkowanego trójwartościowymi jonami lantanowców, tj. iterbu (Yb^{3+}) oraz erbu (Er^{3+}), w którym zaobserwowano efektywną konwersję dwóch niskoenergetycznych fotonów (~1000 nm) na fotony o wyższej energii (540 i 650 nm) [2]. Ze względu na budowę poziomów energetycznych wynikającą z przejść f-f elektronowych up-konwersja jest zjawiskiem charakterystycznym dla jonów lantanowców. Przez ponad trzy dekady zjawisko UC badano i opisywano dla materiałów monokrystalicznych, a także dla mikrometrycznych proszków. Dopiero wraz z rozwojem nanotechnologii na początku lat dziewięćdziesiątych ubiegłego wieku nastąpił gwałtowny wzrost zainteresowania nanoluminoforami domieszkowanymi jonami lantanowców (ang. *Up-converting Nanoparticles*, UCNPs). Nieorganiczne nanomateriały up-konwersyjne zbudowane są z jonów domieszkujących (tzw. *guest*), które stanowią źródło UC, a także z nieorganicznej matrycy (tzw. *host*), której struktura

¹ justyna.czajka@pbs.edu.pl, Zakład Chemii Ogólnej i Nieorganicznej, Wydział Technologii i Inżynierii Chemicznej, Politechnika Bydgoska im. J. i J. Śniadeckich, pbs.edu.pl.

krystalograficzna zapewnia odpowiednie rozmieszczenie centrów luminescencyjnych. Dobór odpowiednich domieszek oraz ich stężeń, a także dobór matrycy pozwala na zsyntetyzowanie nowoczesnych nanomateriałów o właściwościach przystosowanych do konkretnych aplikacji. Aktualnie, nowoczesne materiały up-konwersyjne znajdują zastosowanie w wielu dziedzinach nauki i przemysłu, np. nanofotonice [3], termometrii optycznej [4], medycynie (np. do obrazowania komórek nowotworowych i terapii na poziomie nano), ogniwach słonecznych, czy optoelektronice [5]. Materiały te są także doskonałymi kandydatami do oświetlenia półprzewodnikowego [6], stosowania w białej diodzie elektroluminescencyjnej LED [7] czy jako nowatorskie, precyzyjne materiały działające przeciwdziałaniu podrabianiu [8].

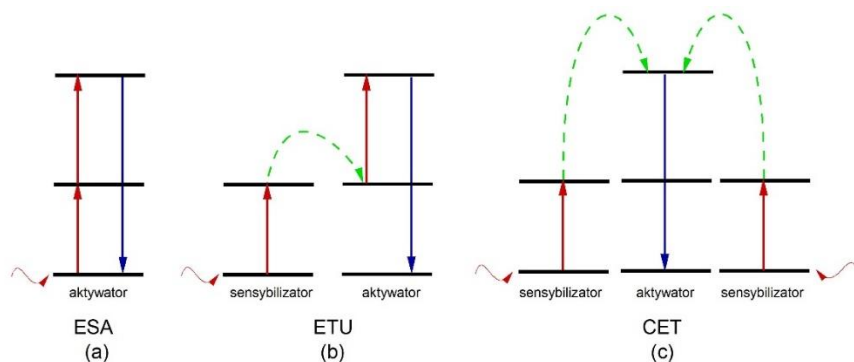
3. Schematy mechanizmów up-konwersji

Różnice między fotoluminescencją a luminescencją up-konwersyjną zaprezentowano na uproszczonym schemacie (rys. 1).



Rysunek 1. Schemat różnic procesów konwencjonalnej luminescencji (a) oraz up-konwersyjnej luminescencji (b) [rysunek własny na podstawie [9]]

Do najważniejszych mechanizmów konwersji w górę opartej na trójwartościowych jonach lantanowców należy absorpcja stanu wzbudzonego (ang. *excited state absorption*, ESA), up-konwersja transferu energii (ang. *energy transfer upconversion*, ETU) oraz kooperacyjny transfer energii (ang. *cooperative energy transfer*, CET). Schematy mechanizmów zaprezentowano na rysunku 2.

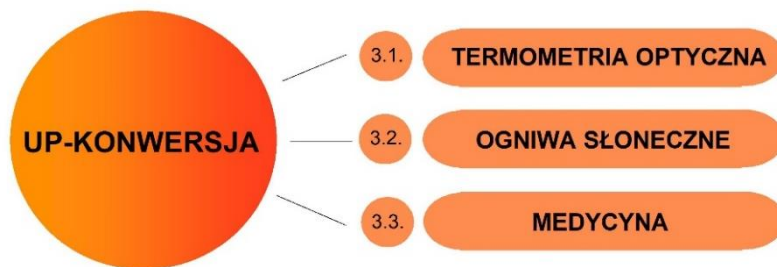


Rysunek 2. Schemat mechanizmów up-konwersji dla trójwartościowych jonów lantanowców, Ln^{3+} : (a) absorpcji stanu wzbudzonego (ESA), (b) up-konwersji transferu energii (ETU), (c) kooperacyjnego transferu energii (CET) [rysunek własny na podstawie [10]]

W mechanizmie absorpcji stanu wzbudzonego (ESA) jeden jon lantanowców(III) kolejno absorbuje dwa fotony, tak że pierwsze wzbudzenie daje metastabilny stan pośredni, który następnie absorbuje drugi foton, aby wyemitować pojedynczy foton o wysokiej energii [9-10]. Mechanizm up-konwersji transferu energii (ETU) zachodzi poprzez sekwencyjną absorpcję fotonów, w układzie, gdzie jeden z dwóch jonów Ln^{3+} pełni funkcję sensybilizatora, a drugi działa jako aktywator, o zbliżonych energiach wzbudzenia oraz będące w niewielkiej odległości od siebie. Mechanizm polega na bezpromienistym przeniesieniu energii od wzbudzonego sensybilizatora do aktywatora, który także znajduje się w stanie wzbudzonym. Jako sensybilizatory najczęściej stosuje się trójwartościowe jony iterbu ze względu na duży przekrój absorpcji oraz położenie jego przejścia f-f, które leży w okolicach 980 nm (przejście ${}^2\text{F}_{7/2} \rightarrow {}^2\text{F}_{5/2}$), co wpływa na ułatwienie transferu energii do najczęściej stosowanych jonów aktywujących, jakimi są erb, holm (Ho^{3+}) i tul (Tm^{3+}) [11]. W trzecim mechanizmie up-konwersji, czyli kooperacyjnym transferze energii (CET), dwa sensybilizatory indywidualnie absorbują fotony o niskiej energii i jednocześnie przekazują swoją energię do aktywatora tak, by wspólnie wzbudzić go w celu późniejszej emisji fotonu o konwersji do góry [10, 12].

4. Wybrane zastosowania up-konwersyjnych nanomateriałów: termometria optyczna, ogniwa słoneczne, medycyna

Nowoczesne materiały up-konwersyjne ze względu na możliwość zaprojektowania pożądanych właściwości strukturalnych, morfologicznych czy luminescencyjnych mają ogromny potencjał w wybranych dziedzinach (rys. 3), takich jak termometria optyczna, ogniwa słoneczne, medycyna (obrazowanie tkanek głębokich, terapia nowotworowa) czy w wielu innych prężnie rozwijających się gałęziach nauki i przemysłu.



Rysunek 3. Wybrane zastosowania up-konwersyjnych nanomateriałów [rysunek własny]

4.1. Termometria optyczna

Zastosowanie nanoluminoforów w termometrii luminescencyjnej pozwala na mapowanie temperatury obiektu z submikrometryczną rozdzielczością przestrzenną, która wykorzystywana jest między innymi w medycynie, biologii komórki, a nawet w przemyśle kosmicznym [13-15]. Wyznaczenie temperatury przeprowadza się na podstawie analizy sygnału luminescencji [12], który zależny jest od temperatury. Najczęściej dokonuje się tego na podstawie natężenia, czasu życia i kształtu pasma [16]. M.S. Pudovkin z zespołem pracował nad termometrami optycznymi opartymi o nanoluminofory typu $\text{Nd}^{3+}, \text{Yb}^{3+}:\text{YF}_3$ badanymi w zakresie temperatur od 80 do 320 K. Materiały luminescencyjne zsyntetyzowano metodą współstrącania z następną obróbką hydrotermiczną (180°C przez 30 h) i wyżarzaniem w próżni w 500°C w czasie 5 h. Otrzymane

luminoformy to nanocząstki o rombowej strukturze krystalicznej oraz o średnicy 48 ± 1 nm [13, 17]. Przyjęto, że efektywny czas życia luminescencji (τ_{eff}) ${}^4F_{3/2}$ neodymu (Nd^{3+}) (przejście ${}^4F_{3/2} \rightarrow {}^4I_{9/2}$, $\lambda_{\text{em}} = 866$ nm) stanowi parametr zależny od temperatury. Ustalono, że τ_{eff} maleje wraz ze spadkiem temperatury w zakresie 80-260 K, natomiast zależność dla zakresu 260-320 K jest mniej wyraźna. Dodatkowo przeprowadzono doświadczenia z zależną od temperatury dyfrakcją promieniowania rentgenowskiego (XRD), dla której zaobserwowano przesunięcie w kierunku mniejszej liczby kątów wraz ze wzrostem temperatury. Zjawisko uzasadniono rozszerzaniem się temperatury sieci krystalicznej [13].

Nanotermometria może zapewnić stabilną, wydajną i nieinwazyjną detekcję temperatury. Jednak konieczność przeprowadzania skomplikowanych syntez, a także czułość na niskie temperatury nadal mocno ogranicza zastosowanie nanometrycznych sond. Z tym problemem zmierzył się Jun Du [18], który wraz z grupą badawczą w swojej publikacji opisał syntezę nowatorskiego up-konwersyjnego nanomateriału na bazie bizmutu z użyciem łatwej i bardzo szybkiej, bo jednodominutowej metody współstrącania. W swojej pracy szczegółowo opisali optymalizację stosunku domieszkowania jonu sensybilizatora Yb^{3+} i jonu aktywatora Er^{3+} oraz wpływ zastosowanego syntetycznego rozpuszczalnika na zwiększenie krystaliczności nanomateriałów i poprawienie intensywności up-konwersji. Podsumowując, uzyskano najlepszą intensywność up-konwersji dla zawartości domieszki jonów Yb^{3+} i jonów Er^{3+} odpowiednio 10% i 2% oraz poprawiono intensywność fluorescencji ponad sto razy po zmianie rozpuszczalnika z glikolu etylenowego (EG), na dimetylosulfotlenek (DMSO). Następnie przeprowadzono badania nad zachowaniem nanomateriału w zakresie termometrii optycznej, tym samym potwierdzając działanie reagowania na temperaturę. Efektem badań było skonstruowanie termometru ratiometrycznego przy użyciu technologii FIR [18].

Według najnowszych doniesień literaturowych nanocząstki typu rdzeń-powłoka $\text{LiErF}_4@ \text{LiYF}_4$ wykazujące termoczulą up-konwersję zaprezentowane na łamach czasopisma „Nanoletters” przez Long Yan [19], wraz z zespołem badawczym, są idealnym kandydatem do aplikacji w bezkontaktowej nanotermometrii. W publikacji zademonstrowano model koncepcyjny celem uzyskania ultraczułej, aktywowanej termicznie up-konwersji w nanocząstkach typu rdzeń-powłoka podsieci erbowej. Dzięki inteligentnej kontroli dynamiki up-konwersji, w której pośredniczy migracja energii jonów erbu, skutkująca wysoką termoczulnością do $5,27\% \text{ K}^{-1}$ przy użyciu niepołączonych termicznie pasm czerwonej i zielonej emisji, która jest znacznie większa od uzyskanej z emisji sprzężonych termicznie, doprowadzono do poprawy konwersji w górę. Zaprezentowany model cechuje się precyzyjną dostrajalną intensywnością luminescencji, kolorami emisji i czasami zaniku, przez co zwiększa się możliwość projektowania nanosystemów multipleksujących [19].

4.2. Ogniwa słoneczne

Jak powszechnie wiadomo paliwa kopalne stanowią aktualnie najbardziej eksploatowane źródło energii. W związku z wysokimi kosztami ich pozyskiwania oraz niebezpiecznymi skutkami, jakie niosą dla ekologii, nieustannie trwają poszukiwania nad uzyskaniem globalnego źródła energii, które rozwiąże problemy energetyczne. Wykorzystanie energii słonecznej, wodnej czy wiatrowej uważane jest za bardziej niezawodną i czystsza formę odnawialnych źródeł energii [20]. Najczęściej ogniwa słoneczne klasy-

fikuje się jako fotowoltaiczne [10, 21], termofotowoltaiczne [22] lub nanofotonowe ogniwa termofotowoltaiczne [23]. Wykorzystaniem up-konwersyjnych luminoforów do zwiększenia pojemności ogniw słonecznych jako pierwszy w 1996 roku zajął się Gibart [24]. W swojej pracy połączył ceramikę współdomieszkowaną jonami iterbu i erbu z ultracienkim ogniwem słonecznym GaAs [10, 22]. W latach 90. prace nad nanostrukturami domieszkowanymi jonami lantanowców wyznaczyły trendy w rozwoju współczesnej nauki o pierwiastkach ziem rzadkich. Najczęściej stosowane są matryce fluorkowe, typu NaYF_4 , NaGdF_4 , LiYF_4 , YF_3 lub tlenkowe domieszkowane jonami iterbu lub erbu. Projektując komponenty strukturalne ogniw słonecznych, należy zoptymalizować np. grubość warstw [25], chropowatość powierzchni granicznych warstw i siatki dyfrakcyjnej [26], czy rodzaj materiałów użytych w ogniwie. Jako sensybilizator najczęściej stosuje się jony iterbu Yb^{3+} ze względu na duży przekrój absorpcji oraz położenie jego przejścia f-f (w okolicach 980 nm), co wpływa na transfer energii do powszechnie stosowanych jonów aktywujących, takich jak Er^{3+} [11], Ho^{3+} i Tm^{3+} [10].

W najnowszym artykule X. Su z zespołem zaproponowali hybrydową metapowierzchnię, która zapewnia skuteczny sposób na osiągnięcie wysoce wydajnej emisji konwersji w górę. W swojej pracy zaproponowali jej wykorzystanie między innymi w ogniwach słonecznych poprzez zastosowanie sensybilizatora Yb^{3+} w UCNPs, wykorzystując ich znaczną absorpcję – ok. 980 nm. Według ich badań pasmo absorpcji plazmonicznej można precyzyjnie kontrolować, tak aby jednocześnie dopasować zarówno długość fali wzbudzenia, jak i emisji nanocząstek konwertujących w górę z domieszką lantanowców (UCNPs). Efektem badań jest zwiększenie up-konwersji teoretycznie o 4-6 rzędów wielkości. Dodatkowo ustalono, że współczynnik wzmocnienia emisji jest wrażliwy na parametry strukturalne oraz współczynnik załamania otaczającego środowiska [27].

4.3. Medycyna

Dane z ostatnich lat wskazują na wzrost liczby osób dotkniętych chorobą nowotworową. Światowa Organizacja Zdrowia (WHO, *World Health Organization*) w raporcie z 2019 roku podała, że tylko w roku 2018 zmarło aż 9,6 mln osób [28, 29]. WHO podaje, że rak jest główną przyczyną zgonów na całym świecie, odpowiadając za prawie 10 milionów zgonów w 2020 roku, czyli odpowiada za prawie co szósty zgon [30]. Dramatyczne spekulacje dotyczące zdrowia na całym świecie przewidują wzrost zachorowań na nowotwory nawet do około 20 milionów nowych pacjentów w 2025 roku [31]. Naukowcy na całym świecie zmagają się między innymi z problemem żywotności komórek po zastosowanym leczeniu.

Zagadnieniem nanomateriałów up-konwersyjnych pod kątem ich zastosowań w fotomedycynie zajął się Lee G. wraz z zespołem badawczym poprzez wykorzystanie ich fotostabilności, multifunkcjonalności oraz biokompatybilności do aplikacji biofotonicznych w głębokich tkankach [32]. Zaprezentowano przegląd najnowszych nieorganicznych i organicznych materiałów UC do zastosowań fotomedycznych, takich jak: terapia fototermiczna (PTT), fotodynamiczna (PDT), chemioterapia wyzwalana światłem, terapia genowa, neuromodulacja za pośrednictwem NIR, fotochemiczne wiązania tkanek (PTB) oraz immunoterapia multimodalna [32].

Badania oparte na użyciu UCNPs do poprawy terapii nowotworowej z minimalizacją negatywnego wpływu leku przeciwnowotworowego na prawidłową tkankę omówiono w artykule przeglądowym opublikowanym w *Advanced Powder Technology* [33]. Autorzy wskazują na jedno z najbardziej wydajnych podejść do przezwyciężenia środków ubocznych terapii, polegające na wytworzeniu biomateriałów w postaci mikro- lub nanokapsulek. W konwersji w górę, istnieje możliwość utworzenia kompleksu typu rdzeń-powłoka (ang. *core-shell*) w celu uzyskania lepszej wydajności. W tej metodzie lek może zostać zamknięty w rdzeniu o wielkości nano lub mikro. Ze względu na wrażliwą na bodźce wewnętrzne i zewnętrzne (pH, różnica temperatur) strukturę powłoki, może zostać ona uszkodzona, co prowadzi ostatecznie do uwolnienia leku z rdzenia do docelowego miejsca [33].

Natomiast w pracy P. Kowalik zaadaptowano nanocząstki NaYF_4 domieszkowane jonami lantanowców, Yb^{3+} i Tm^{3+} o średnicy około 22 nm i bardzo wąskim rozkładzie wielkości, pokryte cienką powłoką SiO_2 (~3 nm), tj. $\text{NaYF}_4:20\% \text{Yb}, 0.2\% \text{Tm}@ \text{SiO}_2$. Nanocząstki ze względu na ich właściwości konwertujące (konwersja światła NIR do światła widzialnego), mogą być stosowane do wykrywania komórek rakowych – 4T1, (zwłaszcza ze względu na emisję 808 nm) bez niepożądanego tła organicznego. Natomiast reaktywne formy tlenu (ang. *reactive oxygen species*, ROS) są generowane przez emisję w górę konwersji jonów Tm/Yb (światło niebieskie i UV). W rezultacie otrzymane wyniki badań terapeutycznych wskazują na duży spadek żywotności komórek po leczeniu [29].

Nowe bionanokompleksy $\text{NaYF}_4:\text{Yb}^{3+}/\text{Er}^{3+}@ \text{silica-TPGS}$ o znacznej biokompatybilności zapewniające silną emisję czerwonej up-konwersji przy bardzo niskiej mocy wzbudzenia, co może stanowić idealną nanocząstkę do znakowania i wykrywania w zastosowaniach biomedycznych opisano w „*Journal of Physical Chemistry B*” [34]. Nanocząstki typu $\text{NaYF}_4:\text{Yb}^{3+}/\text{Er}^{3+}$ otrzymano metodą hydrotermalną, dodatkowo pokryto powłoką krzemionkową zmodyfikowaną metodą Stöbera i sfunkcjonalizowano za pomocą bursztynianu D- α -tokoferylo-glikolu polietylenowego 1000 (TPGS). Stwierdzono, że zsyntetyzowane bionanokompleksy mogą silnie znakować komórki raka piersi MCF-7 do eksperymentów *in vitro* wykrywanych przez mikroskop fluorescencyjny, ale nie może wykryć zdrowych komórek HEK-293A w eksperymentach *in vitro* [34].

Poza diagnostyką nowotworową również medycyna precyzyjna wymaga opracowania nowatorskich rozwiązań, na przykład sond diagnostycznych o doskonałej granicy wykrywalności oraz szerokim zakresie analizy liniowej ze względu na indywidualną rozbieżność w stężeniach choroby. Grupa badawcza Zhou J. skupiła swoje badania na up-konwersyjnej nanosondzie dostosowanej do ultraczułego monitorowania wydychanego tlenu azotu do wskazania klinicznego przebiegu astmy [35].

5. Wnioski/Podsumowanie

Poszukiwanie nowych źródeł energii, dramatycznie rosnący wzrost zachorowań na nowotwory czy rozwój nauki np. w kierunku nanotermometrii skłania naukowców z całego świata do prowadzenia zaawansowanych badań nad nowymi materiałami luminescencyjnymi. Ciągły rozwój nanotechnologii czy dążenie do miniaturyzacji świadczą o konieczności łączenia wielu ścieżek nauki. A. Ghazy w pracy opublikowanej w „*Solar Energy Materials and Solar Cells*” podaje, że w okresie 2017-2020 liczba

artykułów naukowych i referatów dotyczących tematyki materiałów up-konwersyjnych utrzymuje się w stałym tempie ok. 2000 publikacji rocznie [10]. Świadczy to o rosnącym zapotrzebowaniu oraz trendzie w projektowaniu nowoczesnych up-konwersyjnych materiałów z możliwością zaprojektowania ich właściwości morfologicznych, strukturalnych i luminescencyjnych przystosowanych do konkretnych aplikacji. Na podstawie przytoczonych artykułów można wnioskować, że badania nad nowoczesnymi nanomateriałami up-konwersyjnymi, opracowaniem nowych metod ich otrzymywania, np. podczas jednonminutowej syntezy [18] oraz określeniem ich możliwości aplikacyjnych stanowią przyszłość rozwoju nauki i współczesnych technologii. Przykładowo, w medycynie nanocząstki $\text{NaYF}_4:20\% \text{Yb}, 0.2\% \text{Tm}@\text{SiO}_2$ [29] i bionanokompleksy $\text{NaYF}_4:\text{Yb}^{3+}/\text{Er}^{3+}@\text{silica-TPGS}$ [34] rzucają nowe światło w badaniach terapeutycznych. Natomiast hybrydowe metapowierzchnie zapewniające osiągnięcie wysoce wydajnej emisji up-konwersji mają potencjał do aplikacji w ogniwach słonecznych [27].

Literatura

1. Bloembergen N., *Solid state infrared quantum counters*, Physical Review Letters, 2 (3), 1959, s. 84-85.
2. Auzel F., *Compteur quantique par transfert d'energie de Yb^{3+} a Tm^{3+} dans un tungstate mixte et dans un verre germinate*, Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences Serie B, 263.14, 1966, s. 819.
3. Habibia M., Bagheria P., Ghazyania N., Zare-Behtashb H., Heydari E., *3D printed optofluidic biosensor: $\text{NaYF}_4:\text{Yb}^{3+}, \text{Er}^{3+}$ upconversion nano-emitters for temperature sensing*, Sensors and Actuators A, 326, 2021, s. 112734.
4. Gopal R., Manam J., *$\text{SrWO}_4:\text{Er}^{3+}$; an efficient green phosphor for LED and optical thermometry applications*, Journal of Materials Science: Materials in Electronics, 33(37), 2022, s. 21746-21761.
5. Sugumaran P.J., Zhang J., Zhang Y., *Synthesis of stable core-shell perovskite based nano-heterostructures*, Journal of Colloid and Interface Science, 628, 2022, s. 121-130.
6. Kumar A., Manam J., *Observation of up conversion/down conversion luminescence and structural analysis of $\text{La}_2\text{Zr}_2\text{O}_7: \text{Pr}^{3+}$ nano phosphors*, Materials Science in Semiconductor Processing, 148, 2022, s. 106828.
7. Li Q., Yang Y., Que W., Lian T., *Size- and Morphology-Dependent Auger Recombination in CsPbBr_3 Perovskite Two-Dimensional Nanoplatelets and One-Dimensional Nanorods*, Nano Letters 19, 2019, s. 5620-5627.
8. Chen J., Wang S., Linabc J., Chen D., *$\text{CsRe}_2\text{F}_7@\text{glass}$ nanocomposites with efficient up-/down-conversion luminescence: from in situ nanocrystallization synthesis to multi-functional applications*, Journal of Materials Chemistry C, 6, 2018, s. 12364-12370.
9. Zhou J., Liu Q., Feng W., Sun Y., Li F., *Upconversion luminescent materials: Advances and applications*, Chemical Reviews, 115(1), 2015, s. 395-465.
10. Ghazy A., Safdar M., Lastusaari M., Savin H., Karppinen M., *Advances in upconversion enhanced solar cell performance*, Solar Energy Materials and Solar Cells, 230, 2021, s. 111234, <https://doi.org/10.1016/j.solmat.2021.111234>.
11. Huang F., Liu X., Ma Y., Kang S., Hu L., Chen D., *Origin of near to middle infrared luminescence and energy transfer process of $\text{Er}^{3+}/\text{Yb}^{3+}$ co-doped fluorotellurite glasses under different excitations*, Scientific Reports, 5, 2015, s. 5-10.
12. Safdar M., Ghazy A., Lastusaari M., Karppinen M., *Lanthanide-based inorganic-organic hybrid materials for photon-upconversion*, Journal of Materials Chemistry C, 8(21), 2020, s. 6946-6965.
13. Pudovkin M.S., Ginkel A.K., Morozov O.A., Kiiamov A.G., Kuznetsov M.D., *Highly-sensitive lifetime optical thermometers based on $\text{Nd}^{3+}, \text{Yb}^{3+}:\text{YF}_3$ phosphors*, Journal of

- Luminescence, 2022, 249 (September 2021), s. 119037, <https://doi.org/10.1016/j.jlumin.2022.119037>.
14. Piñol R., Brites C.D.S., Silva N.J., Carlos L.D., Millán A., *Nanoscale Thermometry for Hyperthermia Applications*, *Nanomaterials for Magnetic and Optical Hyperthermia Applications*, 2018, s. 139-172.
 15. Jaque D., Vetrone F., *Luminescence nanothermometry*, *Nanoscale*, 15(4), 2012, s. 4301-4326.
 16. Pudovkin M.S., Koryakovtseva D.A., Lukinova E.V., Korableva S.L., Khusnutdinova R.S., Kiiamov A.G., Nizamutdinov A.S., Semashko V.V., *Luminescence Nanothermometry Based on Pr³⁺: LaF₃ Single Core and Pr³⁺: LaF₃/LaF₃ Core/Shell Nanoparticles*, *Advances in Materials Science and Engineering*, 2019, s. 2618307, <https://doi.org/10.1155/2019/2618307>.
 17. Mei Chee Tan, Kumar G.A., Riman R.E., Brik M.G., Brown E., Hommeric U., *Synthesis and optical properties of infrared-emitting YF₃:NdYF₃:Nd nanoparticles*, *Journal of Applied Physics*, 106, 2009, s. 63118, <https://doi.org/10.1063/1.3168442>.
 18. Du J., Liu J., Chen Y., Zhao Y., Li Y., Miao Y., *Lanthanide-doped bismuth-based nanophosphors for ratiometric upconversion optical thermometry*, *RSC Advances*, 12(14), 2022, s. 8743-8749.
 19. Yan L., Huang J., An Z., Zhang Q., Zhou B., *Activating Ultrahigh Thermo-responsive Upconversion in an Erbium Sublattice for Nanothermometry and Information Security*, *Nano Letters*, 22(17), 2022, s. 7042-7048.
 20. Kumari R., Vinayak K.S., Kumar D., *Synthesis Techniques for rare Earth doped up-conversion Nano-materials for Solar cells – A brief Review*, *IOP Conference Series: Earth and Environmental Science*, 889(1), 2021, s. 12057.
 21. Ojha V., Jansen G., Patané A., La Magna A., Romano V., Nicosia G., *Design and characterization of effective solar cells*, *Energy Systems*, 13(2), 2022, s. 355-382.
 22. Shan S., Chen B., Tian J., Zhou Z., Zhang Y., *Improving the performance of steam power cycle through thermo-photovoltaic device: A novel combined system and thermodynamic analysis*, *Energy Conversion and Management*, 269, 2022, s. 116139.
 23. Lenert A., Bierman D.M., Nam Y., Chan W.R., Celanović I., Soljačić M., Wanga E.N., *A nano-photon solar thermophotovoltaic device*, *Nature Nanotechnology*, 9(2), 2014, s. 126-130.
 24. Gibart P., Auzel F., Guillaume J.-C., Zahraman K., *Below Band-Gap IR Response of Substrate-Free GaAs Solar Cells Using Two-Photon Up-Conversion*, *Japanese Journal of Applied Physics*, 35, 1996, s. 4401.
 25. Kim J., Park C., Pawar S.M., Inamdar A.I., Jo Y., Han J., Hong J., Park Y.S., Kim D.Y., Jung W., Kim H., Im H., *Optimization of sputtered ZnS buffer for Cu₂ZnSnS₄ thin film solar cells*, *Thin Solid Films*, 566, 2014, s. 88-92.
 26. Lin A., Phillips J., *Optimization of random diffraction gratings in thin-film solar cells using genetic algorithms*, *Solar Energy Materials and Solar Cells*, 92(12), 2008, s. 1689-1696.
 27. Su X., Bao Q., Zhan S., Liu Y., Nie G., Liu Q., Wu X., *Extraordinary upconversion enhancement in hybrid metasurface by simultaneous excitation/emission band match*, *Journal of Alloys and Compounds*, 925, 2022, s. 166761, <https://doi.org/10.1016/j.jallcom.2022.166761>.
 28. World Health Organization 2019 International Agency for Research on Cancer, *Cancer fact sheets*, *The Global Cancer Observatory* (<https://gco.iarc.fr/>).
 29. Kowalik P., Kamińska I., Fron C.K., Borodziuk A., Duda M., Wojciechowski T., Sobczak K., Kalinowska D., Klepka M.T., Sikora B., *The ROS-generating photosensitizer-free NaYF₄:Yb,Tm@SiO₂ upconverting nanoparticles for photodynamic therapy application*, *Nanotechnology*, 32, 2021, s. 475101.
 30. World Health Organization 3 February 2022, <https://www.who.int/news-room/fact-sheets/detail/cancer>.
 31. Ferlay J., Soerjomataram I., Dikshit R., Eser S., Mathers C., Rebelo M., Parkin D.M., Forman D., Bra F., *Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012*, *International Journal of Cancer*, 136(5), 2015, s. 359-386.

32. Lee G., Choi H.E., Hong S.H., Choi M., Han D.W., Lee J., Kim K.S., Hahn S.K., *Upconversion nanomaterials and delivery systems for smart photonic medicines and healthcare devices*, *Advanced Drug Delivery Reviews*, 188, 2022, s. 114419.
33. Jafari M., Rezanpour A., *Upconversion nano-particles from synthesis to cancer treatment: A review*, *Advanced Powder Technology*, 30, 2019, s. 1731-1753.
34. Thu Huong T., Thi Phuong H., Thi Vinh L., Thi Khuyen H., Thi Thao D., Dac Tuyen L., Kim Anh T., Quoc Minh L., *Upconversion NaYF₄:Yb³⁺/Er³⁺@silica-TPGS Bio-Nano Complexes: Synthesis, Characterization, and in Vitro Tests for Labeling Cancer Cells*, *Journal of Physical Chemistry B*, 125(34), 2021, s. 9768-9775.
35. Zhou J., Wu R., Fu X., Wu J., Mei Q., *Ratio-Adjustable Upconversion Luminescence Nanoprobe for Ultrasensitive In Vitro Diagnostics*, *Analytical Chemistry*, 93, s. 9299-9303.

Charakterystyka spektroskopowa oraz właściwości aplikacyjne nowych up-konwersyjnych nanomateriałów nieorganicznych

Streszczenie

Przez ponad trzydzieści lat up-konwersję, która jest procesem optycznym polegającym na konwersji fotonów o niższej energii w fotony o wyższej energii, opisywano dla mikrometrycznych materiałów. Nagły rozwój nanotechnologii na początku lat 90. ubiegłego wieku doprowadził do wzrostu zainteresowania nanomateriałami, w tym nanoluminoforami domieszkowanymi jonami lantanowców. Światowy trend projektowania nowoczesnych up-konwersyjnych materiałów stanowi punkt wyjścia do pogłębienia aktualnej wiedzy będącym założeniem pracy. Najnowsze doniesienia literaturowe potwierdzają szerokie spektrum zastosowań nowych nanomateriałów, przykładowo w nowych źródłach energii, nanotermometrii czy medycynie. Przytoczone artykuły zawierają obszerną charakterystykę spektroskopową, strukturalną i morfologiczną up-konwersyjnych nanomateriałów, ukazując ich wpływ na właściwości aplikacyjne.

Słowa kluczowe: nanomateriały, up-konwersja, lantanowce, pierwiastki ziem rzadkich

Spectroscopic characteristics and application possibilities of new up-conversion inorganic nanomaterials

Abstract

For over thirty years, up-conversion, an optical process involving the conversion of lower-energy photons into higher-energy photons, has been described for micrometric materials. In the early 90s, the *rapid development* of nanotechnology increased interest in nanomaterials, especially lanthanide ion-doped nanophosphors. The global trend of designing modern up-conversion materials is the starting point for expanding the current knowledge, which is the assumption of this work. The latest literature reports a wide range of applications of new nanomaterials, for example, in new energy sources, nanothermometry, and medicine. The cited articles contain extensive spectroscopic, structural and morphological characteristics of up-conversion nanomaterials, showing their influence on application properties.

Keywords: nanomaterials, up-conversion, lanthanides, rare-earth elements

Heterogeniczne katalizatory renowe

1. Wprowadzenie

Katalizatory heterogeniczne są o wiele częściej stosowane w różnych procesach przemysłowych w porównaniu z katalizatorami homogenicznymi, ze względu na swoje unikatowe właściwości, takie jak: łatwość separacji katalizatora od produktu, możliwość regeneracji katalizatora oraz jego powtórnego użycia w procesie chemicznym. Heterogeniczne układy katalityczne zwykle charakteryzują się wyższą stabilnością w ekstremalnych warunkach pracy (np. temperatura, ciśnienie) niż ich analogii homogeniczne [1-4]. Co ważne w reakcjach z ich udziałem nie jest konieczne stosowanie rozpuszczalników, ponieważ reakcja katalityczna zachodzi na granicy faz. Katalizatory heterogeniczne występują najczęściej w postaci ciała stałego, z tego względu określane są mianem katalizatorów kontaktowych, a układy fazowe tworzone z ich udziałem w procesie katalitycznym stanowią: ciało stałe (katalizator) – ciecz lub ciało stałe (katalizator) – gaz (najczęściej spotykany układy fazowy) [1-4]. Reakcja katalityczna prowadzona w obecności katalizatora heterogenicznego jest procesem składającym się z kilku etapów, obejmujących przede wszystkim: (1) adsorpcję, a następnie chemisorpcję reagentów na powierzchni katalizatora; (2) reakcję na powierzchni katalizatora; (3) desorpcję otrzymanych produktów reakcji z powierzchni katalizatora [1-4].

Katalizatory heterogeniczne charakteryzują się niejednorodną strukturą powierzchni, tylko wybrane fragmenty (miejsca) na powierzchni katalizatora, zwane centrami aktywnymi wykazują aktywność katalityczną w prowadzonych procesach chemicznych. Centrami aktywnymi zlokalizowanymi na powierzchni katalizatora są pojedyncze atomy lub grupy atomów, które są odpowiedzialne za utworzenie wiązania z cząstką reagenta w procesie tworzenia produktu przejściowego bądź kompleksu aktywnego [1]. Ze względu na charakter centra aktywne można podzielić na centra kwasowe i zasadowe (typu Brönsteda i Lewisa) oraz centra redoksove [1-4].

Istotnym elementem każdego katalizatora heterogenicznego jest nośnik, którego rolę może pełnić: tlenek glinu, glinokrzemiany, węgiel aktywny, zeolity czy różne

¹ joanna.malarz@imn.lukasiewicz.gov.pl, Centrum Hydroelektrometalurgii, Sieć Badawcza Łukasiewicz – Instytut Metali Nieżelaznych, www.imn.gliwice.pl.

² katarzyna.leszczyńska-sejda@imn.lukasiewicz.gov.pl, Centrum Hydroelektrometalurgii, Sieć Badawcza Łukasiewicz – Instytut Metali Nieżelaznych, www.imn.gliwice.pl.

³ grzegorz.benke@imn.lukasiewicz.gov.pl, Centrum Hydroelektrometalurgii, Sieć Badawcza Łukasiewicz – Instytut Metali Nieżelaznych, www.imn.gliwice.pl.

⁴ karolina.goc@imn.lukasiewicz.gov.pl, Centrum Hydroelektrometalurgii, Sieć Badawcza Łukasiewicz – Instytut Metali Nieżelaznych, www.imn.gliwice.pl.

⁵ karolina.pianowska@imn.lukasiewicz.gov.pl, Centrum Hydroelektrometalurgii, Sieć Badawcza Łukasiewicz – Instytut Metali Nieżelaznych, www.imn.gliwice.pl.

⁶ julita.sztandera@imn.lukasiewicz.gov.pl, Centrum Hydroelektrometalurgii, Sieć Badawcza Łukasiewicz – Instytut Metali Nieżelaznych, www.imn.gliwice.pl.

⁷ michal.ochmanski@imn.lukasiewicz.gov.pl, Centrum Hydroelektrometalurgii, Sieć Badawcza Łukasiewicz – Instytut Metali Nieżelaznych, www.imn.gliwice.pl.

formy krzemionki. W procesach prowadzonych na skalę przemysłową najczęściej jako nośnik stosowany jest tlenek glinu, przykładem jest tutaj proces reformingu katalitycznego. Niekiedy celem zwiększenia aktywności i selektywności uzyskanego układu katalitycznego konieczna jest jego dodatkowa aktywacja poprzez wprowadzenie tzw. promotora reakcji. Rolę promotora mogą pełnić metale, tlenki metali, kwasy nieorganiczne/organiczne [1-4].

Katalizatory heterogeniczne znacznie częściej w porównaniu z ich analogami homogenicznymi ulegają efektowi zatrucia, co przejawia się częściową lub całkowitą utratą aktywności katalitycznej. Efekt zatrucia układu katalitycznego następuje w wyniku trwałej adsorpcji na powierzchni katalizatora trucizny bądź poprzez reakcję trucizny z katalizatorem, co prowadzi do utworzenia nieaktywnego katalitycznie kompleksu. Do najbardziej znanych, trucizn heterogenicznych układów katalitycznych należą: H_2S , N_2 , siarczki organiczne i nieorganiczne, związki As, P, Pb oraz Hg [1-4].

2. Ren

Ren jest niezwykle cennym metalem rzadkim, którego zawartość w skorupie ziemskiej szacowana jest na $7 \cdot 10^{-8}\%$ (0,7 ppm) [5, 6]. Największe pokłady tego metalu znajdują się w Chile, USA, Chinach, Kazachstanie, Armenii oraz Rosji. Głównym producentem tego metalu ze źródeł pierwotnych jest Chile obok USA i Polski [5, 7]. Ren otrzymywany jest jako produkt uboczny przy przetwarzaniu koncentratów miedzi oraz molibdenu. Z kolei wiodącymi producentami renu ze źródeł wtórnych są USA i Niemcy obok takich krajów, jak: Kanada, Estonia, Francja, Japonia, Polska i Rosja. Ren na rynkach globalnych sprzedawany jest najczęściej w postaci: (1) peletek, (2) brykietów, (3) proszku, (4) renianu(VII) amonu, (5) kwasu renowego(VII), (6) proszków stopowych [8-11]. W Polsce ren oferowany jest w postaci renu metalicznego (peletek) oraz renianu(VII) amonu o czystości katalitycznej przez firmę KGHM Metraco S.A. [8]. Drugim krajowym producentem tego metalu jest firma Innovator Sp. z o.o. mająca w swoim portfolio takie produkty, jak: renian(VII) amonu, kwas renowy(VII), renian(VII) niklu(II), renian(VII) kobaltu(II), renian(VII) srebra, czy stopy oraz proszki stopowe na bazie Co lub Ni [9]. Renian(VII) amonu o czystości katalitycznej wytwarzany jest zarówno z odpadów pochodzących z przerobu koncentratów miedzi (przez firmę KGHM Metraco S.A.), jak również z odpadów pochodzących z recyklingu głównie złomu superstopów (przez firmę Innovator Sp. z o.o.) [8, 9].

Ren jest srebro-szarym metalem, który charakteryzuje się wysoką temperaturą topnienia ($3180^\circ C$) oraz wrzenia ($5870^\circ C$), dużą gęstością ($2,1 \text{ g/dm}^3$), twardością oraz odpornością na ścieranie [5, 6]. Metal ten jest odporny na korozję, roztwarza się w kwasach utleniających, nie jest podatny na działanie kwasu chlorowodorowego oraz fluorowego.

2.1. Zastosowanie renu

Ren ze względu na swoje specyficzne właściwości znajduje szerokie spektrum zastosowań w wielu gałęziach przemysłu. Metal ten stosowany jest głównie jako dodatek do superstopów na bazie niklu, które mogą zawierać od 1 do nawet 7% renu, a wykorzystywane są one przede wszystkim do budowy łopatek turbin będących elementami silników odrzutowych, turbin gazowych czy osłon pojazdów kosmicznych. Z kolei

stopy wolframowo-renowe stosowane są w lampach rentgenowskich oraz do produkcji termoelementów [6].

Ren znajduje również zastosowanie w katalizie, w szczególności w katalizie heterogenicznej. Heterogeniczny układ katalityczny $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$, który charakteryzuje się wysoką aktywnością i selektywnością w procesie metatezy olefin, wykorzystywany jest między innymi do otrzymywania propylenu z etenu i 2-butenu. Warto w tym miejscu wspomnieć, że metateza olefin jest ważną reakcją dla kilku sektorów przemysłu chemicznego (syntezy organicznej, chemii polimerów, petrochemii). Bimetaliczne katalizatory Pt-Re osadzone na stałych nośnikach są stosowane w przemyśle petrochemicznym, np. w procesie reformingu surowej benzyny. Metal ten pełni rolę promotora w katalizatorach srebrnych stosowanych w procesach epoksydacji (bezpośredniego utleniania) etylenu do tlenku etylenu. Ren dodawany jest również do katalizatorów kobaltowych oraz żelazowych wykorzystywanych w technologii GTL (ang. *gas to liquid*), umożliwiających produkcję czysto spalających się paliw diesla, LPG, oleju bazowego i nafty z gazu ziemnego poprzez konwersję węglowodorów gazowych do węglowodorów o dłuższych łańcuchach, które są ciekłe w warunkach otoczenia. Jest to metoda konwersji gazu ziemnego do ciekłych produktów o wyższej jakości, które w innym przypadku byłyby produkowane z ropy naftowej [6].

Izotopy ^{186}Re i ^{188}Re są przedstawicielami dwóch niezwykle ważnych dla medycyny nuklearnej radionuklidów. Izotopy te znajdują zastosowanie w diagnostyce oraz terapeutycznej medycynie nuklearnej do leczenia guzów nowotworowych [6].

3. Heterogeniczne katalizatory renowe

Katalizatory renowe, jak już wspomniano we wcześniejszym podrozdziale, stosowane są w wielu procesach prowadzonych na skalę przemysłową, np. w reformingu benzyny ciężkiej. Znajdują one również zastosowanie w różnego typu reakcjach chemicznych, np. reakcji metatezy olefin, reakcji hydrodeoksygenacji (HDO).

3.1. Monometaliczne katalizatory – $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$

Tlenek renu(VII) osadzony na γ -tlenku glinu ($\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$) jest bardzo dobrze znanym w literaturze heterogenicznym katalizatorem metatezy olefin [12-16]. Typowy katalizator otrzymywany jest poprzez impregnację tlenku glinu charakteryzującego się rozwiniętą powierzchnią ($180\text{-}200\text{ m}^2/\text{g}$) wodnym roztworem renianu(VII) bądź kwasu renowego(VII). Po odparowaniu wody prekursor katalizatora renowego suszony jest w temperaturze 110°C , a w kolejnym kroku poddawany jest procesowi kalcynacji w zakresie temperatury wynoszącym $500\text{-}550^\circ\text{C}$. Podczas procesu kalcynacji powstaje, krystaliczny tlenek renu(VII) (Re_2O_7), o temperaturze topnienia wynoszącej 297°C , który jest równomiernie rozproszony na nośniku. Co istotne, Re_2O_7 jest lotny i łatwo sublimuje poniżej temperatury topnienia; w związku z tym w wyniku procesu kalcynacji następuje jego częściowa utrata. Uzyskany w przedstawiony sposób heterogeniczny katalizator $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ nie zawiera w swej sferze koordynacyjnej ugrupowania alkilidenu, dlatego po wprowadzeniu alkeny, następuje redukcja Re^{7+} i w ten sposób tworzony jest kompleks metal-alken, inicjujący proces metatezy. Czasami stosowane są związki metaloorganiczne, które pełnią funkcję promotora etapu inicjacji katalizatora bądź zwiększają aktywność układu katalitycznego.

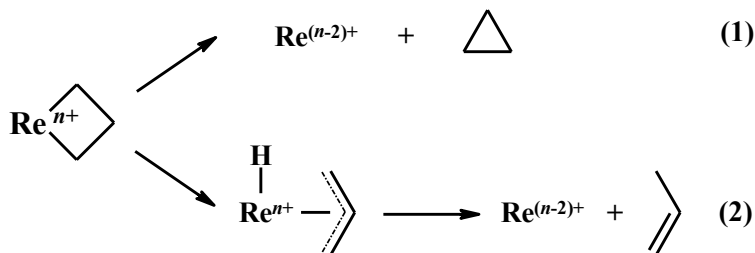
W reakcjach katalizowanych przez katalizatory heterogeniczne niezwykle istotna jest struktura powierzchni katalizatora, ponieważ to właśnie na niej zachodzą wszelkie procesy. Struktura powierzchni układu $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ charakteryzowana jest przez zastosowanie szerokiego spektrum technik, w tym metod spektroskopowych, także w połączeniu technik eksperymentalnych z teoretycznymi obliczeniami metodą DFT [17-19].

Wiadomo, że układ $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ charakteryzuje się wysoką aktywnością katalityczną oraz selektywnością w niskiej temperaturze (20-100°C), co bez wątplenia jest jego zaletą. W przypadku metatezy olefin z ugrupowaniami funkcyjnymi konieczne jest więc zastosowanie promotora, którego rolę pełnią związki tetraalkilocyny (SnR_4 ; R = metyl, etyl lub butyl) lub tetraalkiloolowiu (PbR_4 ; R = etyl lub butyl) [20-22]. Katalizatory $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ ze związkami tetraalkilocyny są znacznie bardziej aktywne, w porównaniu do układów bez promotora, również w procesie metatezy niepodstawionych alkenów, ale niestety szybciej ulegają dezaktywacji. Aktywność takich układów katalitycznych można przywrócić częściowo poprzez ich kalcynowanie w strumieniu powietrza oraz ponowne dodanie promotora. Jednak, po każdym cyklu regeneracji, następuje stopniowa utrata aktywności katalizatora w wyniku akumulacji cyny (najprawdopodobniej w postaci SnO_2) na powierzchni katalizatora [23].

Jak twierdzili autorzy jednej z prac, popierając wysuniętą tezę własnymi badaniami, prawdopodobnie aktywność katalizatora $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ uzależniona jest również od maksymalnej ilości Re_2O_7 , jaka może zostać „zakotwiczona” na powierzchni Al_2O_3 [24]. Co zostało wyjaśnione na podstawie badań powierzchniowych grup OH^- tlenku glinu przy użyciu spektroskopii FTIR. Ostatnio, Panpranot wraz ze współpracownikami wykazali, że aktywne są również katalizatory $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ zawierające zaledwie 4% mas. Re. Taki układ został zastosowany w metatezie prop-2-enu i etylenu [25], przy czym użyta maksymalna ilość Re_2O_7 (14-18% mas. Re_2O_7) uzależniona była od rodzaju tlenku glinu, od jego powierzchni właściwej oraz objętości porów [26, 27].

Ponadto, zaobserwowano, że przy małej zawartości procentowej renu, aktywność układu katalitycznego $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ można zwiększyć poprzez stosowanie wyższej temperatury kalcynacji. Optymalna temperatura procesu zależy więc od zawartości renu w katalizatorze i zawiera się w zakresie temperatury wynoszącym od 825 do 925°C [28]. Co ciekawe, aktywność układu $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ zwiększana jest poprzez działanie na Al_2O_3 kwasem chlorowodorowym lub fluorowodorowym, który reaguje z zasadowymi grupami hydroksylowymi znajdującymi się na powierzchni tlenku glinu [29]. Dodatkowo aktywność katalizatorów zwiększana może być poprzez: (1) dodawanie tlenków metali przejściowych (MoO_3 , WO_3 lub V_2O_5) [30]; (2) stosowanie podłoża mieszanego ($\text{SiO}_2\text{-Al}_2\text{O}_3$, $\text{Al}_2\text{O}_3\text{-B}_2\text{O}_3$) [20, 31, 32]; (3) wykorzystywanie fosforanu(V) glinu [33]. Na przykład stosując podłoże mieszane, zwiększana jest wydajność heterogenicznego katalizatora, w wyniku wyższej kwasowości i większej ilości centrów kwasowych Brönsteda, niż to ma miejsce w przypadku typowego katalizatora Re. Stosując katalizatory renowe, spotykamy się z wcześniej już wspomnianym problemem ich dezaktywacji. Jedną z przyczyn tego zjawiska jest zatrucie katalizatora przez związki polarne, np. wodę obecną w stosowanych reagentach [15, 34, 35]. Pozostałe to: (1) redukcja centrów metalicznych w czasie reakcji poniżej optymalnego stopnia utlenienia; (2) blokowanie centrów aktywnych poprzez adsorpcję ubocznych produktów polimeryzacji na powierzchni katalizatora; (3) redukcyjna eliminacja pośredniego produktu – kompleksu metalocyklobutanu [15, 35]. Dezaktywacja katalizatora renowego jest nie unikniona,

prawdopodobnie główną przyczyną jest redukcyjna eliminacja kompleksu metalo-cyklobutanu prowadząca do utworzenia cyklopropanu (ścieżka 1) lub β -eliminacja do alkenu (ścieżka 2) – proponowany mechanizm przedstawiono na rysunku 1 [15, 16, 35].



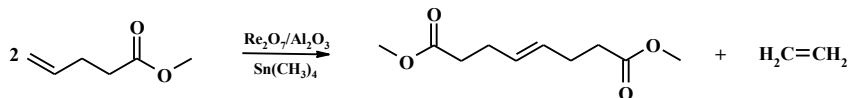
Rysunek 1. Redukcyjna eliminacja metylocyklobutanu, prowadzi do otrzymania: (1) cyklopropanu; (2) alkenu
Źródło: [15, 16]

Pomimo stosunkowo szybkiej dezaktywacji układ katalityczny $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$, po dodatkowych modyfikacjach stosowany był często ze względu: (1) na możliwość prowadzenie reakcji w niskiej temperaturze, (2) łatwy sposób regeneracji zużytego katalizatora.

Katalizatory renowe wykorzystywane są przede wszystkim w procesie metatezy gazowych i/lub ciekłych alkenów oraz acyklicznych alkenów z ugrupowaniami funkcyjnymi. Na przykład, w jednej z pierwszych prac grupa japońskich badaczy opisała możliwość wykorzystania układu $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ w reakcji metatezy heks-1-enu. Proces prowadzony był w reaktorze przepływowym ze stałym złożem, pod ciśnieniem atmosferycznym, w temperaturze $31,5^\circ\text{C}$ [36]. Następnie, Mol wraz ze współpracownikami przeprowadzili reakcję metatezy propenu, stosując katalizator $\text{Re}_2\text{O}_7/\text{SiO}_2 \cdot \text{Al}_2\text{O}_3$, w mikro-katalitycznym reaktorze przepływowym ze złożem stałym, tym samym uzyskując wyższą aktywność katalityczną w porównaniu do układu $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ [20]. Z kolei Commereuc i inni wykorzystali $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ w metatezie butenu i etylenu oraz odwrotnie propylenu, co ciekawe proces prowadzony był w fazie ciekłej w reaktorach integralnych, pod ciśnieniem, w temperaturze zbliżonej do otoczenia [34]. W kolejnej pracy Kawai i jego współpracownicy opisali reakcję *self*-metatezy oraz *cross*-metatezy liniowych alkenów ($\text{RCH} = \text{CHR}'$; przy czym $\text{R}' < \text{R}$; $\text{R}' = \text{H}, \text{CH}_3, \text{C}_2\text{H}_5, \text{C}_3\text{H}_7$) zawierających od pięciu do dziewięciu atomów węgla, wobec układu $\text{Re}_2\text{O}_7/\text{CsNO}_3/\text{Al}_2\text{O}_3$, prowadzoną w fazie ciekłej, w szklanym reaktorze przepływowym, ze stałym złożem katalizatora, pod ciśnieniem atmosferycznym [37]. $\text{Re}_2\text{O}_7/\text{CsNO}_3/\text{Al}_2\text{O}_3$ zastosowano także: w metatezie α, ω -dienów [38], 1,4-alkadienów [39] czy metylo-1-alkenów [40]. Układy $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ otrzymane w wyniku użycia mezoporowatego Al_2O_3 , stosowano w procesie metatezy heks-1-enu, prowadząc reakcje w naczyniu typu Schlenk, w temperaturze 40°C , w atmosferze osuszonego azotu [41]. Ponadto, katalizatory $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ uzyskane na bazie makro-mezoporowatego oraz mezoporowatego tlenku glinu wykorzystywane były w procesie metatezy but-1-enu i but-2-enu [42, 43]. W ostatnim czasie układ $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ stosowano w reakcji metatezy terminalnych alkenów (pent-1-en, heks-1-en, hept-1-en, okt-1-en) prowadzonej w reaktorze przepływowym oraz wsadowym [44].

Początkowo wykonane badania przez Nakamura oraz współpracowników wskazały, iż sam układ $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ nie jest aktywny w procesie metatezy związków polarnych [45]. Jako pierwszy Verkuijlen wraz ze współpracownikami opisał metatezę nienasy-

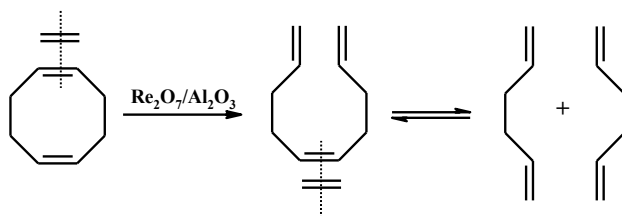
conych estrów kwasów tłuszczowych, na przykładzie pent-4-enu metylu wobec $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ z dodatkiem niewielkiej ilości tetrametylocyny, jako promotora (rysunek 2), reakcję prowadzona była w szklanym reaktorze wsadowym, w temperaturze 50°C , przez 1 h [46].



Rysunek 2. Metateza pent-4-enu metylu z udziałem $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ i $\text{Sn}(\text{CH}_3)_4$ [46]
Źródło: Opracowanie własne w oparciu [46]

Układy $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3\text{-SnR}_4$ (R = metyl, etyl lub butyl) stosowane były w wielu reakcjach metatezy nienasyconych estrów [21, 47, 48], w procesie metatezy etylenu (etenolizy) z alkenami często z ugrupowaniem funkcyjnym do otrzymywania terminalnych alkenów [21]. Katalizator $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3\text{-SnR}_4$ wykorzystywany był także w reakcji metatezy olefin zawierających chlor czy brom [21, 49].

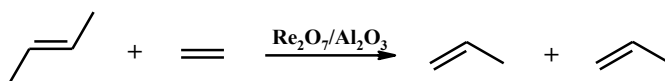
Układ katalityczny $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ stosowany był również w procesach technologicznych, prowadzonych na większą skalę. Pod koniec lat 80. ubiegłego stulecia, firma Schell opracowała proces o nazwie FEAST. W wyniku dysproporcjonowania (etenolizy) cyklooktadienów wobec $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$, w łagodnych warunkach (temperatura: $0\text{-}20^\circ\text{C}$; ciśnienie: 1-2 bar) otrzymywano α,ω -diolefiny [12, 50, 51]. Na przykład poprzez reakcję cyklookta-1,5-dienu z etenem uzyskiwano heksa-1,5-dienu – co zaprezentowano na rysunku 3.



Rysunek 3. Synteza heksa-1,5-dienu *via* etenoliza cyklookta-1,5-dienu [50]
Źródło: Opracowanie własne w oparciu [50]

Proces FEAST skomercjalizowano, w 1987 roku otwarto fabrykę w Etang de Berre (Francja), o zdolności produkcyjnej diolefin wynoszącej 3000 ton/rok. Jednak po kilku latach została zamknięta ze względu na brak zapotrzebowania rynkowego na produkty [12, 51].

Kilka lat później, Instytut Francais du Petrole (IFP) wspólnie z Chinese Petroleum Corporation (Tajwan) opracował proces wytwarzania propenu o nazwie META-4®. But-2-enu reaguje z etylenem w fazie ciekłej, w obecności układu $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$, w niskiej temperaturze 35°C , pod ciśnieniem 60 bar (rys. 4) [12, 51-53].



Rysunek 4. Metateza but-2-enu z etenem wobec $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ [12]
Źródło: Opracowanie własne w oparciu [12]

Niewątpliwą zaletą tego katalizatora jest jego ciągła regeneracja. Co więcej, w latach 1988-1990, w Tajwanie działała instalacja pilotażowa o zdolności produkcji propenu wynoszącej 15 kg/godzinę [51]. Proces ten do tej pory nie został skomercjalizowany, głównie ze względu na wysokie koszty katalizatora oraz konieczność użycia reagentów o wysokiej czystości.

Wykorzystując katalizatory heterogeniczne w reakcji metatezy, można otrzymywać związki typu *fine chemicals*, np. feromony owadów. Tradycyjna droga syntezy tej grupy związków jest sekwencją kilku, a nawet kilkunastu etapowych reakcji. Zastosowanie metatezy olefin pozwala na skrócenie syntezy do kilku etapów, co wiąże się z wyższą wydajnością oraz czystością produktów końcowych. Na przykład w skali laboratoryjnej *cis*-9-trikozen (feromon płciowy muchy domowej) otrzymywany był w krzyżowej metatezie 9-oktadecenu z 2-heksadecenem wobec katalizatora $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ [54]. Także, główny składnik perfum muskon uzyskiwany był w reakcji metatezy, w obecności właśnie $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ [16]. Z literatury wiadomo, że związki typu *fine chemicals* otrzymywane są w wyniku metatezy funkcjonalizowanych alkenów. I tak, metateza estrów nienasyconych kwasów tłuszczowych przy użyciu układu $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ jest dogodną i wysoce selektywną drogą syntezy nienasyconych diestrów, które mogą być stosowane do produkcji wartościowych produktów, np. związków makrocyklicznych.

3.2. Bimetaliczne katalizatory – Pt-Re/ Al_2O_3

Bimetaliczne katalizatory Pt-Re osadzone na γ -tlenku glinu ($\text{Pt-Re}/\text{Al}_2\text{O}_3$) są stosowane w procesie reformingu benzyny ciężkiej, do produkcji benzyny o wyższej liczbie oktanowej. W procesie reformingu węglowodorów nasyconych zachodzą głównie reakcje izomeryzacji, dehydrogenacji, dehydrocyklizacji oraz hydrokrakingu, celem wytworzenia węglowodorów aromatycznych o wyższej liczbie oktanowej [6, 55-58]. Proces reformingu katalitycznego prowadzony jest w temperaturze 450-520°C i przy ciśnieniu 7-68 atm [57]. Bimetaliczne układy $\text{Pt-Re}/\text{Al}_2\text{O}_3$ to katalizatory dwufunkcyjne, w których funkcje metaliczne odpowiedzialne są za reakcje dehydrogenacji oraz hydrogenolizy, a funkcje kwasowe kontrolują reakcje izomeryzacji oraz cyklizacji [55-58]. Katalizatory platynowo-renowe charakteryzują się dużą stabilnością, tolerancją na zanieczyszczenia surowcowe oraz przede wszystkim łatwością w regeneracji. Układy katalityczne $\text{Pt-Re}/\text{Al}_2\text{O}_3$ stosowane w procesach przemysłowych zawierają: 0,2-0,3% mas. Pt; 0,20-0,75% mas. Re; 0,60-1,3% mas. chlorków; ~98% mas. $\gamma\text{-Al}_2\text{O}_3$ [57]. Poniżej w tabeli 1. przedstawiono skład chemiczny oraz właściwości bimetalicznego katalizatora reformingu $\text{Pt-Re}/\text{Al}_2\text{O}_3$ o nazwie handlowej R-62 [6].

Tabela 1. Skład chemiczny oraz właściwości bimetalicznego katalizatora reformingu katalitycznego $\text{Pt-Re}/\text{Al}_2\text{O}_3$ [6]

KATALIZATOR – R-62	
kształt cząstek	sferyczne
średnica cząstek, mm	1,6
gęstość usypowa, lb/ft ³	44,3
zawartość Pt, % mas.	0,22
zawartość Re, % mas.	0,44
zawartość Cl, % mas.	1,0-1,1

Źródło: [6].

Nośnikiem katalizatora Pt-Re/Al₂O₃ jest γ -tlenek glinu o rozwiniętej powierzchni właściwej, mieszczącej się w zakresie od 180 do 230 m²/g, który dodatkowo modyfikowany jest jonami chlorkowymi [57]. Działanie katalizatora uzależnione jest od stężenia jonów chlorkowych oraz wzajemnego oddziaływania powierzchniowych grup hydroksylowych tlenku glinu z jonami chlorkowymi, które odpowiadają za kwasową funkcję izomeryzacji katalizatorów reformingu [57]. Zatem wprowadzenie jonów chlorkowych ma za zadanie zagwarantować odpowiednią kwasowość nośnika katalizatora.

Katalizatory Pt-Re/Al₂O₃ wytwarzane są przez impregnację tlenku glinu roztworami wodnymi, związków, takich jak: kwas heksachloroplatynowy(IV), renian(VII) amonu (lub kwas renowy(VII)), kwas chlorowodorowy [57]. Po usunięciu wody prekursor układu katalitycznego w pierwszej kolejności poddawany jest kalcynacji w atmosferze tlenu, a następnie redukcji w atmosferze wodoru. Podobnie jak wcześniej wspomniane katalizatory Re₂O₇/Al₂O₃, również te badane są kompleksowo pod kątem morfologii powierzchni za pomocą różnych metod spektralnych [57, 59].

Ren dodany do katalizatora platynowego zwiększa jego stabilność, działa jako promotor i nie katalizuje żadnej innej specyficznej reakcji. Obecność tego metalu w układzie katalitycznym zmniejsza szybkość jego dezaktywacji, umożliwiając wydłużenie czasu pracy układu przy stosunkowo niskim ciśnieniu. Dezaktywacja katalizatorów reformingu spowodowana jest osadzaniem się koksu na powierzchni katalizatora, polega to na sukcesywnej fragmentacji i odwodornieniu węglowodorów oraz reorganizacji form węgla do grafitowych warstw węgla [55].

Katalizatory bimetaliczne Pt-Re/Al₂O₃ są także podatne na zatrucie. Główną trucizną katalizatorów reformingu jest siarka (w postaci siarki organicznej występującej w takich związkach, jak: tiofeny, alkioltiofeny, merkaptany), której adsorpcja na powierzchni katalizatora Pt-Re w niewielkiej ilości jest korzystna. Jednak nadmierne siarkowanie wpływa silnie na ich aktywność katalityczną. Toksyczność siarki wynika zarówno z blokowania miejsc aktywnych, jak i modyfikacji właściwości elektronowych tych miejsc. Jest ona powszechnie znaną trucizną funkcji hydrogenacji/dehydrogenacji katalizatorów reformingu katalitycznego. Zazwyczaj przy stężeniu siarki poniżej 1 ppm, modyfikowane są i zatrywane centra aktywne katalizatora. Zatrucie katalizatora siarką może spowodować tymczasową bądź trwałą jego dezaktywację w procesie reformingu. Stopień zatrucia układu katalitycznego uzależniony jest od czasu kontaktu katalizatora i stężenie siarki w przetwarzanej benzynie ciężkiej. Stężenie siarki w benzynie ciężkiej kierowanej do procesu reformingu katalitycznego powinno mieścić się w zakresie od 0,2 do 0,5 ppm [57]. Dlatego ważne jest, aby stosować odpowiednie warunki obróbki wstępnej wsadu ciężkiej benzyny, aby uniknąć silnego zatrucia katalizatora. Surowce stosowane w procesach reformingu katalitycznego zwykle są poddawane, najpierw hydrowodorafinacji w celu usunięcia zanieczyszczeń [57-59].

Badania nad układami katalitycznymi typu Pt-Re/Al₂O₃ prowadzone są od wielu lat. Bazowy bimetaliczny katalizator Pt-Re/Al₂O₃ poddawany był różnym modyfikacjom poprzez wprowadzenie promotora reakcji, tj.: metalu – np.: Ge [60-62], Sn [63-66], In [67, 68], Ga [69]; kwasu nieorganicznego lub organicznego [70, 71] (tab. 2). Zabiegi te mają wpłynąć na poprawę jego właściwości katalitycznych w procesie reformingu, a mianowicie zwiększyć aktywność oraz selektywność, ale głównie wydłużyć jego żywotność. Przykłady otrzymanych układów katalitycznych (bi- oraz trimetalicznych) w wyniku modyfikacji bimetalicznego katalizatora bazowego Pt-Re/Al₂O₃ przedstawiono

w tabeli 2. Podano tam także warunki, w jakich przeprowadzane zostały reakcje reformingu *n*-alkanów wobec nowych układów katalitycznych. Próby z zastosowaniem tych katalizatorów wykonywane były w reaktorach przepływowych ze złożem stałym.

Trimetaliczne katalizatory uzyskiwane są w wyniku sukcesywnej impregnacji bądź metodą współimpregnacji odpowiednich prekursorów metali na tlenku glinu. Sposób otrzymywania tych trójskładnikowych układów katalitycznych wpływa na właściwości zarówno funkcji kwasowej, jak i metalicznej katalizatora reformingu, a w konsekwencji na jego aktywność, selektywność oraz odporność na dezaktywację.

Na przykład, przy otrzymywaniu katalizatorów Pt-Re-Ge/Al₂O₃, kolejność wprowadzania prekursorów metali na tlenek glinu w dużym stopniu wpływa na właściwości i wydajność tych układów katalitycznych w procesie reformingu *n*-oktanu [60]. Dodanie Pt przed Re oraz Ge powoduje, że metale te są redukowane oddzielnie i w efekcie uzyskuje się katalizator charakteryzujący się wysoką aktywnością w reakcji dehydrogenacji oraz mniejszą w reakcji hydrogenolizy. Z drugiej strony, impregnacja Ge przed Pt i Re prowadzi do uzyskania katalizatora o niskiej aktywności w reakcji dehydrogenacji i wysokiej aktywności w reakcji hydrogenolizy, co jest konsekwencją silnych oddziaływań między metalami. Natomiast po wprowadzenie Re na tlenek glinu przed Pt i Ge otrzymuje się katalizatora o właściwościach pośrednich. Zatem kolejność wprowadzania prekursorów metali wpływa na funkcje metaliczne układu katalitycznego. Jak twierdzą autorzy pracy, Ge oraz Re zmieniają kwasowość nośnika w sąsiedztwie metali, efektem tego są różnice w selektywności w stosunku do związków aromatycznych i izomerów parafin [60]. Katalizatory Re-Pt-Ge/Al₂O₃ charakteryzują się większą selektywnością wobec rozgałęzionych izoparafin, natomiast przy użyciu katalizatora Ge-Re-Pt/Al₂O₃ można uzyskać reformat bogaty w związki aromatyczne [60].

Tabela 2. Katalizatory stosowane w badaniach nad procesem katalitycznego reformingu *n*-alkanów

Lp.	Katalizator/ Literatura	Promotor	Proces reformingu	Warunki procesu reformingu katalitycznego
1	Pt-Re-Ge/Al ₂ O ₃ [60]	Ge (GeCl ₄)	<i>n</i> -oktan	T = 450°C, P = 1 atm, t = 6 h
2	Pt-Re-Ge/Al ₂ O ₃ [61]	0,1% mas. Ge (GeCl ₄)	<i>n</i> -heptan	T = 450°C, P = 0,1 MPa, WHSV = 7 h ⁻¹ , H ₂ / <i>n</i> -C ₇ H ₁₆ = 4 mol/mol
3	Pt-Re-Ge/Al ₂ O ₃ [62]	0,1-2,0% mas. Ge (GeCl ₄ w H ₂ O, NH ₃ (aq), 0,2 M HCl)	<i>n</i> -heptan	T = 500°C, P = 0,5 MPa H ₂ / <i>n</i> -C ₇ H ₁₆ = 7,5 mol/mol
4	Pt-Re-Sn/Al ₂ O ₃ [64]	0,3% mas. Sn (SnCl ₂)	<i>n</i> -oktan, <i>n</i> -pentan	T = 500°C, P = 1 atm, WHSV = 1,8 h ⁻¹ (<i>n</i> -C ₈ H ₁₈), H ₂ / <i>n</i> -C ₈ H ₁₈ = 10 mol/mol, WHSV = 4,5 h ⁻¹ (<i>n</i> -C ₅ H ₁₂), H ₂ / <i>n</i> -C ₅ H ₁₂ = 6 mol/mol
5	Pt-Re-Sn/Al ₂ O ₃ [66]	0,06-0,32% mas. Sn (SnCl ₂ ·2H ₂ O)	<i>n</i> -oktan	T = 480°C, P = 10 bar, t = 6 h WHSV = 3,3 h ⁻¹ , H ₂ / <i>n</i> -C ₈ H ₁₈ = 4,5
6	Pt-Re-In/Al ₂ O ₃ [67, 68]	0,1-0,3% mas. In (In(NO ₃) ₃)	<i>n</i> -heptan	T = 500°C, P = 0,1 MPa, WHSV = 4 h ⁻¹ , H ₂ / <i>n</i> -C ₇ H ₁₆ = 6

7	Pt-Re-Ga/Al ₂ O ₃ [69]	0,1-0,3% mas. Ga (GaNO ₃)	<i>n</i> -heptan	T = 500°C, P = 0,1 MPa, WHSV = 4 h ⁻¹ , H ₂ / <i>n</i> -C ₇ H ₁₆ = 6
8	Pt-Re/Al ₂ O ₃ [70]	modyfikacja nośnika – HF (0,4-2,0% mas. F)	<i>n</i> -heptan	T = 460-520°C, P = 1,5 MPa LHSV = 8-14 h ⁻¹ , H ₂ / <i>n</i> -C ₇ H ₁₆ = 5 mol/mol
9	Pt-Re/Al ₂ O ₃ [71]	modyfikacja nośnika – CH ₃ COOH (0,83-17,4 mol/dm ³)	<i>n</i> -heptan	T = 460-520°C, P = 1,0 MPa LHSV = 8,2-14 h ⁻¹ , H ₂ / <i>n</i> -C ₇ H ₁₆ = 5 mol/mol

LHSV (ang. *Liquid Hourly Space Velocity*) – objętość gazu wchodzącego do reaktora przez masę katalizatora na godzinę, WHSV (ang. *Weight Hourly Space Velocity*) – masa gazu wchodzącego do reaktora przez masę katalizatora na godzinę.

Źródło: Opracowanie własne w oparciu [60-62, 64, 66-71].

Rodzaj oraz ilość wprowadzonego promotora do katalizatora platynowo-renowego również ma wpływ na charakter funkcji kwasowej, jak i metalicznej katalizatora, a w konsekwencji tego na przebieg procesu katalitycznego reformingu. Dodatek do układu bazowego Pt-Re/Al₂O₃ takich metali, jak Ge lub Sn ma korzystny wpływ na funkcje kwasowe katalizatora, czego przejawem jest zwiększenie się jego aktywności w reakcji izomeryzacji [60, 61, 64, 66]. Z kolei wprowadzenie In lub Ga wpływa na zmniejszenie się jego kwasowości oraz zahamowanie funkcji metalicznej, tj. aktywności w reakcji dehydrogenacji oraz hydrogenolizy [68-69]. W przypadku wprowadzenia In efekt zmniejszania się kwasowości katalizatora pogłębia się wraz ze wzrostem jego zawartości w układzie katalitycznym [68].

3.3. Pozostałe katalizatory renowe

Heterogeniczne katalizatory renowe poza procesem katalitycznego reformingu oraz metatezy olefin stosowane są także w różnych reakcjach chemicznych, np. reakcji hydrodeoksygenacji (HDO) czy reakcji deoksydehydratacji (DODH). W tabeli 3 zestawiono przykłady heterogenicznych katalizatorów renowych znajdujących zastosowanie w ww. reakcjach chemicznych. Jak wskazuje analiza danych zawartych w tabeli 3, reakcja HDO oraz DODH może zachodzić wobec mono- lub bimetalicznych katalizatorów renowych. Katalizatory wykorzystywane w tych reakcjach otrzymywane są, podobnie jak już wcześniej omówione układy katalityczne, w procesie kalcynacji bądź w wyniku połączenia procesu kalcynacji oraz redukcji.

Tabela 3. Katalizatory renowe stosowane w reakcjach hydrodeoksygenacji (HDO) oraz deoksydehydratacji (DODH)

Lp.	Katalizatory/Literatura	Substrat/ Typ reakcji	Najbardziej aktywny układ katalityczny
1	Re/SiO ₂ , Re/TiO ₂ , Re/ZrO ₂ , Re/CeO ₂ , ReO _x /TiO ₂ , ReO _x /Al ₂ O ₃ [72]	2-metoksyfenol HDO	Re/SiO ₂ – konwersja 2-metoksyfenolu 98%, otrzymano cykloheksan/benzen z 58% selektywnością
2	Ir-Re/Al ₂ O ₃ , Re/Al ₂ O ₃ [73]	izoeugenol HDO	Ir-Re/Al ₂ O ₃ – otrzymano propylocykloheksan z 99% wydajnością

Heterogeniczne katalizatory renowe

3	Re/SiO ₂ , Ni-Re/SiO ₂ [74]	2-metoksyfenol HDO	Ni-Re/SiO ₂ – konwersja 2-metoksyfenolu 100%, otrzymano cykloheksan z ~95% selektywnością
4	ReO _x /SiO ₂ , ReO _x /Al ₂ O ₃ , ReO _x /TiO ₂ (czynnik redukujący – 2-heksanol) [75]	glicerol DODH	ReO _x /Al ₂ O ₃ – konwersja glicerolu >99%, otrzymano alkohol alilowy z 91% wydajnością
5	ReO _x /C (czynnik redukujący – H ₂) [76]	(+)-winian dietylu DODH	ReO _x /C – otrzymano fumaran dietylu z 95% wydajnością
6	ReO _x -Ru/CeO ₂ , ReO _x -Rh/CeO ₂ , ReO _x -Pt/CeO ₂ , ReO _x -Pd/CeO ₂ , ReO _x -Ir/CeO ₂ , ReO _x -Au/CeO ₂ (czynnik redukujący – H ₂) [77]	glicerol DODH	ReO _x Au/CeO ₂ – konwersja glicerolu >99%, otrzymano alkohol alilowy z 91% wydajnością

Źródło: Opracowanie własne w oparciu [72-77].

W procesie reformingu w fazie wodnej (APR, ang. *Aqueous Phase Reforming*) rozpuszczone w wodzie węglowodory lub oksygenaty reagują w fazie wodnej z cząsteczką wody, prowadząc do otrzymania mieszaniny gazów wzbogaconych w wodór [78, 79]. Proces ten prowadzony jest w niskiej temperaturze 220-270°C, pod wysokim ciśnieniem 30-60 bar, w obecności układu katalitycznego [80]. W reakcji tej również znajdują zastosowanie zarówno mono-, jak i bimetaliczne katalizatory renowe, w szczególności platynowo-renowe. Poniżej przedstawiono przykłady renowych układów katalitycznych wykorzystywanych w procesie APR (tab. 4).

Tabela 4. Katalizatory renowe stosowane w reakcji reformingu w fazie wodnej (APR)

Lp.	Katalizatory	APR	Literatura
1	Pt-Re/C, Re/C (bez i z dodatkiem – KOH)	glicerol	[81]
2	Re/C, Pt-Re/C, Pt-Re/SiO ₂		[82]
3	Rh-Re/C, Pt-Re/C, Re/C		[83]
4	Pt-Re/TiO ₂	ksylitol	[84]
5	Re/TiO ₂ , Ru-Re/TiO ₂ , Rh- Re/TiO ₂ , Ir-Re/TiO ₂ , Pt-Re/TiO ₂	etanol	[85]
6	Pt-Re/CMK-3 (mezoporowaty węgiel)	glikol etylenowy	[86]

Źródło: Opracowanie własne w oparciu [81-86].

4. Podsumowanie

Renowe katalizatory heterogeniczne znajdują zastosowanie w procesach przemysłowych oraz różnego typu reakcjach chemicznych prowadzonych głównie w skali laboratoryjnej. Z punktu widzenia katalizy heterogenicznej, najważniejszym układem katalitycznym jest Pt-Re/Al₂O₃, stosowany w procesie katalitycznego reformingu benzyny

ciężkiej. Drugim dosyć istotnym katalizatorem jest $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ ze względu na zastosowanie w powszechnie znanej reakcji metatezy olefin. Tego typu układy otrzymywane są w procesie kalcynacji lub w wyniku połączenia procesu kalcynacji oraz redukcji. Aktywność oraz selektywność w procesie katalitycznym ww. katalizatorów można zmieniać poprzez wprowadzenie promotora. I tak, dodatek tetralkilocyny lub tetralkiloolowiu do układu $\text{Re}_2\text{O}_7/\text{Al}_2\text{O}_3$ ma wpływ na poprawę jego aktywności w procesie metatezy olefin z grupowaniami funkcyjnymi. W przypadku katalizatorów bimetalicznych Pt-Re/ Al_2O_3 dodatek promotora zazwyczaj modyfikuje funkcję metaliczną oraz funkcję kwasową tego układu, co przejawia się zmianą aktywności oraz selektywności w procesie reformingu katalitycznego. Jak już wspomniano, katalizatory heterogeniczne są częściej podatne na zatrucia niż ich analogii homogeniczne, czego efektem jest częściowa bądź całkowita utarta ich aktywności katalitycznej. Jednak z powodzeniem mogą być regenerowane. Katalizatory renowe stosowane są również w reakcji hydrodeoksygenacji (HDO), deoksyhydratacji (DODH) oraz w procesie reformingu w fazie wodnej (APR).

Podsumowując, podczas projektowania nowych układów katalitycznych należy wziąć pod uwagę następujące aspekty:

- sposób zastosowania katalizatora – typ reakcji chemicznej;
- rodzaj i ilość składników wpływających na aktywność potencjalnego katalizatora znajdującego zastosowanie w dedykowanej reakcji chemicznej;
- typ oraz rodzaj podłoża dla potencjalnego katalizatora;
- metodę przygotowania katalizatora (w tym sposób impregnacji podłoża, kolejność dodawania poszczególnych prekursorów metali);
- ilość oraz rodzaj dodawanego promotora reakcji.

Uwagi ogólne i podziękowania

Serdeczne podziękowania dla wszystkich osób, które umożliwiły publikację tego rozdziału i pomogły w jego tworzeniu.

Praca finansowana w ramach:

- projektu NanoRen, pt. „Innowacyjna technologia wytwarzania nanokomponentów Re z odpadów pochodzących z recyklingu”, Działanie 4.1, Poddziałania 4.1.4 Projekty Aplikacyjne Programu Operacyjnego Inteligentny Rozwój 2014-2020, współfinansowanego ze środków Europejskiego Funduszu Rozwoju Regionalnego i NCBR, POIR.04.01.04-00-0008/20;
- Norweskiego Mechanizmu Finansowego 2014-2021 – Small Grant 2020 NOR/SGS//RenMet/0049/2020-00 (11/PE/0146/21), pt. „Innovative hydrometallurgical technologies for the production of rhenium compounds from recycled waste materials for catalysis, electromobility, aviation and defense industry”.

Literatura

1. Grzybowska-Swierkosz B., *Elementy Katalizy Heterogenicznej*, Wydawnictwo Naukowe PWN, Warszawa 1993, ISBN 83-01-10511-9.
2. Rase H.F., *Handbook of Commercial Catalysts. Heterogeneous Catalysts*, CRC Press Taylor & Francis Group, 2000, ISBN 978-1-4822-7536-0.
3. Davis M.E., Davis R.J., *Fundamentals of Chemical Reaction Engineering*, McGraw-Hill, USA, New York 2003, ISBN 0-07-119260-3.

4. Ertl G., Knözinger H., Weitkamp J., *Handbook of Heterogeneous Catalysis*, VCH Wiley Company, 3, 1997, ISBN 9783527619474.
5. Nadler H.G., Starck H.C., *Rhenium and Rhenium Compounds*, in Ullmann's Encyclopedia of Industrial Chemistry, 31, John Wiley & Sons Inc., USA, New York 2005, s. 527-537.
6. *Rhenium: Market Outlook to 2029 Report*, 11th ed., Roskill Information Services Ltd., 2019.
7. Polyak D.E., *Mineral Commodity Summaries 2022 – Rhenium* (usgs.gov), <https://pubs.usgs.gov/periodicals/mcs2022/mcs2022-rhenium.pdf> [data dostępu: 17.10.2022].
8. <https://metraco.pl/produkty-i-uslugi/nadrenian-oraz-ren/> [data dostępu: 17.10.2022].
9. <http://www.innovator.com.pl/ren/> [data dostępu: 17.10.2022].
10. <https://molymet.com/en/productos-servicios/productos-renio/> [data dostępu: 17.10.2022].
11. https://www.heraeus.com/en/hpm/hmp_products_solutions/rhenium/rhenium.html [data dostępu: 17.10.2022].
12. Ivin K.J., Mol J.C., *Olefin Metathesis and Metathesis Polymerization*, Academic Press, 1997, ISBN 0-12-377045-9.
13. Grella K., *Olefin Metathesis: Theory and Practice*, John Wiley & Sons, Inc., Hoboken, NJ, USA 2014, ISBN 978-1-118-20794-9.
14. Grubbs R.H., *Handbook of Metathesis: Catalyst Development*, Wiley-VCH Verlag GmbH, 2003, ISBN 9783527619481.
15. Mol J.C., *Olefin metathesis over supported rhenium oxide catalysts*, Catalysis Today, 51, 1999, s. 289-299.
16. Copéret C., Berkson Z.J., Chan K.W., de Jesus Silva J., Gordon C.P., Pucino M., Zhizhko P.A., *Olefin metathesis: what have we learned about homogeneous and heterogeneous catalysts from surface organometallic chemistry?*, Chemical Science, 12, 2021, 3092-3115.
17. Vicente B.C., Nelson R.C., Moses A.W., Chattopadhyay S., Scott S.L., *Interactions Involving Lewis Acidic Aluminum Sites in Oxide-Supported Perrhenate Catalysts*, The Journal of Physical Chemistry, 115, 2011, s. 9012-9024.
18. Bare S.R., Kelly S.D., Vila F.D., Boldingh E., Karapetrova E., Kas J., Mickelson G.E., Modica F.S., Yang N., Rehr J.J., *Experimental (XAS, STEM, TPR, and XPS) and Theoretical (DFT) Characterization of Supported Rhenium Catalysts*, The Journal of Physical Chemistry, 115, 2011, s. 5740-5755.
19. Lwin S., Keturakis Ch., Handzlik J., Sautet P., Li Y., Frenkel A.I., Wachs I.E., *Surface ReO_x Sites on Al_2O_3 and Their Molecular Structure-Reactivity Relationships for Olefin Metathesis*, ACS Catalysis, 5, 2015, s. 1432-1444.
20. Andreini A., Xiaoding X., Mol J.C., *ACTIVITY OF $Re_2O_7/SiO_2-Al_2O_3$ CATALYSTS FOR PROPENE METATHESIS AND THE INFLUENCE OF ALKYLITIN PROMOTORS*, Applied Catalysis, 27, 1986, s. 31-40.
21. Mol J.C., *Metathesis of Functionalizes Acyclic Olefins*, Journal of Molecular Catalysis, 65, 1991, s. 145-162.
22. Xiaoding X., Mol J.C., *$Re_2O_7/SiO_2-Al_2O_3-SnR_4$ or $-PbR_4$, a Highly Active Catalyst for the Metathesis of Functionalized Alkenes*, Journal of the Chemical Society, Chemical Communications 1, 1985, s. 631-633.
23. Spronk R., Mol J.C., *Regeneration of rhenium-based catalysts for the metathesis of propene*, Applied Catalysis, 76, 1991, s. 143-152.
24. Kapteijn F., Bredt L.H.G., Mol J.C., *Recueil des Travaux Chimiques des Pays-Bas*, 96, 1997, s. 139.
25. Vorakitkanvasin S., Na Ayudhya S.K., Suriye K., Praserttham P., Panpranot J., *Enhanced metathesis activity of low loading Re_2O_7/Al_2O_3 catalysts for propylene production by using aluminum nitrate as Al_2O_3 precursor*, Applied Catalysis A: General, 517, 2016, s. 39-46.

26. Vuurman M.A., Stufkens D.J., Oskam A., Wachs I.E., *Structural determination of surface rhenium oxide on various oxide supports (Al₂O₃, ZrO₂, TiO₂ and SiO₂)*, Journal of Molecular Catalysis, 76, 1992, s. 263-285.
27. Ellison A., Coverdale A.K., Dearing P.F., *Catalytic metathesis of unsaturated esters by alumina supported rhenium oxide*, Applied Catalysis, 8, 1983, s. 109-121.
28. Spronk R., van Veen J.A.R., Mol J.C., *The Effect of Calcination Temperature on the Activity of Re₂O₇/Al₂O₃ Catalysts for the Metathesis Propene*, Journal of Catalysis, 144, 1993, s. 472-483.
29. Hietala J., Root A., Knuutila P., *The Surface Acidity of Pure and Modified Aluminas in Re/Al₂O₃ Metathesis Catalysts as Studied by ¹H MAS NMR Spectroscopy and Its Importance in the Ethenolysis of 1,5-Cyclooctadiene*, Journal of Catalysis, 150, 1994, s. 46-55.
30. Xiaoding X., Imhoff P., van den Aerweg G.C.N., Mol J.C., *Mixed-oxide Catalysts for the Metathesis of Functionalized Alkenes*, Journal of the Chemical Society, Chemical Communications, 1985, s. 273-275.
31. Xiaoding X., Boelhouwer C., Benecke J.I., Vonk D., Mol J.C., *Re₂O₇/Al₂O₃-B₂O₃ Metathesis Catalysts*, Journal of the Chemical Society, Faraday Transactions, 1(82), 1986, s. 1945-1953.
32. Mol J.C., Andreini A., *Activity and selectivity of rhenium-based catalysts for alkene metathesis*, Journal of Molecular Catalysis, 46, 1988, s. 151-156.
33. Sibeijn M., Spronk R., van Veen J.A.R., Mol J.C., *Ir studies OF Re₂O₇ metathesis catalysts supported on alumina and phosphated alumina*, Catalysis Letters, 8, 1991, s. 201-208.
34. Amigues P., Chauvin Y., Commereuc D., Hong C.T., Lai C.C., Liu Y.H., *Methathesis of ethylene-butene mixtures to propylene with rhenium on alumina catalysts*, Journal of Molecular Catalysis, 65, 1991, s. 39-50.
35. Spronk R., Andreini A., Mol J.C., *Deactivation of rhenium-based catalysts for the metathesis of propene*, Journal of Molecular Catalysis, 65, 1991, s. 219-235.
36. Kawai T., Yamazaki Y., Takumura A., *Metathesis of 1-Hexene over Re₂O₇-Al₂O₃ Catalysts in Liquid Phase*, Journal of the Japan Petroleum Institute, 26, 1983, s. 332-338.
37. Kawai T., Goto H., Yamazaki Y., Ishikawa T., *Metathesis of n-alkenes over a CsNO₃-Re₂O₇-Al₂O₃ catalyst*, Journal of Molecular Catalysis, 46, 1988, s. 157-172.
38. Kawai T., Yamazaki Y., Taoka T., Kobayashi K., *Metathesis of α, ω-Dienes over a CsNO₃-Re₂O₇-Al₂O₃ Catalyst in Liquid Phase*, Journal of Catalysis, 89, 1984, s. 452-461.
39. Kawai T., Goto H., Ishikawa T., Yamazaki Y., *Metathesis of 1,4-alkadienes over a CsNO₃-Re₂O₇-Al₂O₃ catalyst*, Journal of Molecular Catalysis, 39, 1987, s. 369-382.
40. Kawai T., Maruoka N., Goke M., Ishikawa T., *Metathesis of methyl-1-alkenes over a CsNO₃-Re₂O₇-Al₂O₃ catalyst*, Journal of Molecular Catalysis, 49, 1989, s. 261-269.
41. Aguado J., Escola J.M., Castro M.C., Paredes B., *Metathesis of 1-hexene over rhenium oxide supported on ordered mesoporous aluminas: Comparison with Re₂O₇/γ-Al₂O₃*, Applied Catalysis A: General, 284, 2005, s. 47-57.
42. Sang L., Chen S.-L., Yuan G., Zheng M., You J., Chen A., Li R., Chen L., *Metathesis of 1-butene and 2-butene to propene over Re₂O₇ supported on macro-mesoporous γ-alumina prepared via a dual template method*, Journal of Natural Gas Chemistry, 21, 2012, s. 105-108.
43. Sang L., Chen S.-L., Yuan G., Zhou Z., Li R., Chen A., Zheng M., You J., *Preparation of mesoporous alumina with large pore size and their supported rhenium oxide catalysts in metathesis of 1-butene and 2-butene to propene*, Journal of Natural Gas Chemistry, 21, 2012, s. 352-359.
44. Kustov L.M., Furman D.B., Barkova A.P., *Metathesis of C5-C8 Terminal Olefins on Re₂O₇/Al₂O₃ Catalysts*, Catalysis Letters, 146, 2016, s. 1033-1039.
45. Nakamura R., Matsumoto S., Echigoya E., *Metathesis of alkenes having functional groups*, Chemistry Letters, 1976, s. 1019-1024.

46. Verkuijlen E., Kapteijn F., Mol J.C., Boelhouwer C., *Heterogeneous metathesis of unsaturated fatty acid esters*, Journal of the Chemical Society, Chemical Communications, 198, 1977, s. 198-199.
47. Ellison A., Coverdale A.K., Dearing P.F., *The metathesis of unsaturated esters over Re_2O_7/Al_2O_3 catalysts*, Journal of Molecular Catalysis, 28, 1985, s. 141-167.
48. Sibeijn M., Mol J.C., *Activity of supported Re_2O_7 catalysts for the metathesis of methyl oleate*, Applied Catalysis, 67, 1990, s. 279-295.
49. Kawai T., Uejima S., Suzuki T., Iyoda T., *Metathesis of halogen-containing olefin over Re_2O_7/Al_2O_3 catalyst promoted with alkylmetal as a cocatalyst*, Journal of Molecular Catalysis A: Chemical, 133, 1998, s. 51-59.
50. Chaumont P., John C.S., *Olefin Disproportionation Technology (FEAST) – a Challenge for Process Development*, Journal of Molecular Catalysis, 46, 1988, s. 317-328.
51. Khosravi E., Szymańska-Buzar T., *Ring Opening Metathesis Polymerisation And Related Chemistry*, Kluwer Academic Publishers, 2000, ISBN 1-4020-0558-X.
52. Chiusoli G.P., Maitlis P.M., *Metal-catalysis in Industrial Organic Processes*, Royal Society of Chemistry, Cambridge 2006, ISBN 978-0854041503.
53. Amigues P., Chauvin Y., Commereuc D., Lai C.C., Liu Y.H., Pan J.M., *Hydrocarbon Process*, 69, 1990, s. 79.
54. Streck R., Naturforsch Z., *Synthese von Insektenlockstoffen an Metathesekatalysatoren, Synthesis of Insect Attractants by Metathesis Catalysts*, Zeitschrift für Naturforschung, 31b, 1976, s. 1256-1264.
55. Xiao J., Puddephatt R.J., *Pt-Re clusters and bimetallic catalysts*, 143, 1995, s. 457-500.
56. Speight J.G., *Handbook of Petroleum Refining, Introduction to refining processes*, Chapter 4, CRC Press, 2016, ISBN 9781315374079.
57. Oyekan S.O., *Catalytic Naphtha Reforming Process*, CRC Press Taylor & Francis Group, 2019, ISBN 9780367780968.
58. Gjervan T., Prestvik R., Holmen A., *Catalytic reforming, Basic Principles in Applied Catalysis*, Ed. M. Baerns, Springer-Verlag Berlin Heidelberg, 2004, eBook ISBN 978-3-662-05981-4.
59. Rønning M., Gjervan T., Prestvik R., Nicholson D.G., Holmen A., *Influence of pretreatment temperature on the bimetallic interactions in Pt-Re/ Al_2O_3 reforming catalysts studied by X-ray absorption spectroscopy*, Journal of Catalysis, 204, 2001, s. 292-304.
60. Carvalho L.S., Conceição K.C.S., Mazzieri V.A., Reyes P., Pieck C.L., do C. Rangel M., *Pt-Re-Ge/ Al_2O_3 catalysts for n-octane reforming: Influence of the order of addition of the metal precursors*, Applied Catalysis A: General, 419-420, 2012, s. 156-163.
61. Mazzieri V.A., Pieck C.L., Vera C.R., Yori J.C., Grau J.M., *Effect of Ge content on the metal and acid properties of Pt-Re-Ge/ Al_2O_3 -Cl catalysts for naphtha reforming*, Applied Catalysis A: General, 353, 2009, s. 93-100.
62. D'Ippolito S.A., Vera C.R., Epron F., Samoila P., Especel C., Marécot P., Gutierrez L.B., Pieck C.L., *Catalytic Properties of Pt-Re/ Al_2O_3 Naphtha-Reforming Catalysts Modified by Germanium Introduced by Redox Reaction at Different pH Values*, Industrial & Engineering Chemistry Research, 3, 2009, s. 3771-3778.
63. Carvalho L.S., Pieck C.L., Rangel M.C., Figoli N.S., Grau J.M., Reyes P., Parera J.M., *Trimetallic naphtha reforming catalysts. I. Properties of the metal function and influence of the order of addition of the metal precursors on Pt-Re-Sn/ γ - Al_2O_3 .Cl*, Applied Catalysis A: General, 269, 2004, s. 91-103.
64. Carvalho L.S., Pieck C.L., Rangel M.C., Figoli N.S., Vera C.R., Parera J.M., *Trimetallic naphtha reforming catalysts II. Properties of the acid function and influence of the order of addition of the metallic precursors on Pt-Re-Sn/ γ - Al_2O_3 .Cl*, Applied Catalysis A: General, 269, 2004, s. 105-116.

65. Pieck C.L., Vera C.R., Parera J.M., Giménez G.N., Serra L.R., Carvalho L.S., Rangel M.C., *Metal dispersion and catalytic activity of trimetallic Pt-Re-Sn/Al₂O₃ naphtha reforming catalysts*, 108, 2005, s. 637-642.
66. Elfghi F.M., Amin N.A.S., *Influence of tin content on the texture properties and catalytic performance of bi-metallic Pt-Re and tri-metallic Pt-Re-Sn catalyst for n-octane reforming*, Reaction Kinetics Mechanism and Catalysis, 114, 2015, s. 229-249.
67. Vicerich M.A., Oportus M., Benitez V.M., Reyes P., Pieck C.L., *Influence of time and temperature on the regeneration of PtReIn/Al₂O₃ naphtha reforming catalysts*, Catalysis Letters, 144, 2014, s. 1178-1187.
68. Benitez V.M., Pieck C.L., *Influence of Indium Content on the Properties of Pt-Re/Al₂O₃ Naphtha Reforming Catalysts*, Catalysis Letters, 136, 2010, s. 45-51.
69. Vicerich M.A., Especel C., Benitez V.M., Epron F., Pieck C.L., *Influence of gallium on the properties of Pt-Re/Al₂O₃ naphtha reforming catalysts*, Applied Catalysis A: General, 407, 2011, s. 49-55.
70. Belopukhov E.A., Kir'yanov D.I., Smolikov M.D., Shkurenok V.A., Belyi A.S., Lavrenov A.V., Kleimenov A.V., Kondrashev D.O., *Investigation of fluorine-promoted Pt-Re/Al₂O₃ catalysts in reforming of n-heptane*, Catalysis Today, 378, 2021, s. 113-118.
71. Tregubenko V.Y., Belyi A.S., *Characterization of Acid-Modified Alumina as a Support for Reforming Catalysts*, Kinetics Catalysis, 61, 2020, s. 130-136.
72. Jeong Y., Park C.W., Park Y.-K., Ha J.-M., Jeong Y., Lee K.-Y., Jae J., *Investigation of the activity and selectivity of supported rhenium catalysts for the hydrodeoxygenation of 2-methoxyphenol*, Catalysis Today, 375, 2021, s. 164-173.
73. Alda-Onggar M., Mäki-Arvela P., Eränen K., Aho A., Hemming J., Paturi P., Peurla M., Lindblad M., Simakova I.L., Dmitry Yu. Murzin, *Hydrodeoxygenation of Isoeugenol over Alumina-Supported Ir, Pt, and Re Catalysts*, ACS Sustainable Chemistry and Engineering, 6, 2018, s. 16205-16218.
74. Park C.W., Kim J.W., Kim H.U., Park Y.-K., Lam S.S., Ha J.-M., Jae J., *Bimetallic Ni-Re catalysts for the efficient hydrodeoxygenation of biomass-derived phenols*, International Journal of Energy Research, 45, 2021, s. 16349-16361.
75. Kon Y., Araque M., Nakashima T., Paul S., Dumeignil F., Katryniok B., *Direct Conversion of Glycerol to Allyl Alcohol Over Alumina-Supported Rhenium Oxide*, Chemistry Select, 2, 2017, s. 9864-9868.
76. Denning A.L., Dang H., Liu Z., Nicholas K.M., Jentoft F.C., *Deoxydehydration of Glycols Catalyzed by Carbon-Supported Perrhenate*, Chem Cat Chem, 5, 2013, s. 1-4.
77. Tazawa S., Ota N., Tamura M., Nakagawa Y., Okumura K., Tomishige K., *Deoxydehydration with Molecular Hydrogen over Ceria-Supported Rhenium Catalyst with Gold Promoter*, ACS Catalysis, 6, 2016, s. 6393-6397.
78. Cortright R.D., Davda R.R., Dumesic J.A., *Hydrogen from catalytic reforming of biomass-derived hydrocarbons in liquid water*, Nature, 418, 2002, s. 964-967.
79. Pipitone G., Zoppi G., Pirone R., Bensaid S., *A critical review on catalyst design for aqueous phase reforming*, International Journal of Hydrogen Energy, 47, 2022, s. 151-180.
80. Davda R.R., Shabaker J.W., Huber G.W., Cortright R.D., Dumesic J.A., *A review of catalytic issues and process conditions for renewable hydrogen and alkanes by aqueous-phase reforming of oxygenated hydrocarbons over supported metal catalysts*, Applied Catalysis B: Environmental, 56, 2005, s. 171-186.
81. King D.L., Zhang L., Xia G., Karim A.M., Heldebrant D.J., Wang X., Peterson T., Wang Y., *Aqueous phase reforming of glycerol for hydrogen production over Pt-Re supported on carbon*, Applied Catalysis B: Environmental, 99, 2010, s. 206-213.
82. Ciftci A., Lighthart D.A.J.M., Sen A.O., Van Hoof A.J.F., Friedrich H., Hensen E.J.M., *Pt-Re synergy in aqueous-phase reforming of glycerol and the water-gas shift reaction*, Journal of Catalysis, 311, 2014, s. 88-101.

83. Ciftci A., Lighthart D.A.J.M., Hensen E.J.M., *Aqueous phase reforming of glycerol over Re-promoted Pt and Rh catalysts*, Green Chemistry, 16, 2014, s. 853-863.
84. Kirilin A.V., Tokarev A.V., Manyar H., Hardacre C., Salmi T., Mikkola J.-P., Murzin D.Yu., *Aqueous phase reforming of xylitol over Pt-Re bimetallic catalyst: Effect of the Re addition*, Catalysis Today, 223, 2014, s. 97-107.
85. Nozawa T., Yoshida A., Hikichi S., Naito S., *Effects of Re addition upon aqueous phase reforming of ethanol over TiO₂ supported Rh and Ir catalysts*, International Journal of Hydrogen Energy, 40, 2015, s. 4129-4140.
86. Kim H.D., Park H.J., Kim T.-W., Jeong K.-E., Chae H.-J., Jeong S.-Y., Lee C.-H., Kim C.-U., *Hydrogen production through the aqueous phase reforming of ethylene glycol over supported Pt-based bimetallic catalysts*, International Journal of Hydrogen Energy, 37, 2012, s. 8310-8317.

Heterogeniczne katalizatory renowe

Streszczenie

Ren stosowany jest głównie jako dodatek do superstopów na bazie niklu oraz wykorzystywany jest do produkcji katalizatorów. Renowe katalizatory heterogeniczne znajdują zastosowanie w procesach przemysłowych oraz w różnych typach reakcji chemicznych. Re₂O₇/Al₂O₃ katalizuje reakcję metatezy olefin. Katalizatory Pt-Re/Al₂O₃ stosowane są w procesie reformingu benzyny ciężkiej do produkcji benzyny o wyższej liczbie oktanowej. Aktywność oraz selektywność ww. układów katalitycznych można modyfikować poprzez wprowadzenie promotora reakcji, tj. metalu, tlenku metalu, kwasu nieorganicznego/organicznego. Katalizatory renowe znajdują zastosowanie w reakcji hydrodeoksygenacji (HDO) i deoksyhydratacji (DODH) oraz w procesie reformingu w fazie wodnej (APR).

Słowa kluczowe: heterogeniczne katalizatory, ren, Pt-Re/Al₂O₃, Re₂O₇/Al₂O₃, kataliza

Heterogeneous rhenium catalysts

Abstract

Rhenium is mainly used as an additive for nickel-based superalloys and is used in the production of catalysts. Rhenium heterogeneous catalysts are used in industrial processes and in various types of chemical reactions. Re₂O₇/Al₂O₃ catalyzes the olefin metathesis reaction. Pt-Re/Al₂O₃ catalysts are used in the reforming of naphtha to produce higher octane number gasoline. The activity and selectivity of the previously mentioned catalytic systems can be modified by introducing a reaction promoter, i.e. metal, metal oxide, inorganic/organic acid. Rhenium catalysts are used in hydrodeoxygenation (HDO) and deoxydehydration (DODH) reactions and in aqueous phase reforming (APR).

Keywords: heterogeneous catalysts, rhenium, Pt-Re/Al₂O₃, Re₂O₇/Al₂O₃, catalysis

Wpływ dodatku odpadowych wiórów stalowych na wartości modułu Young'a ekokompozytów cementowo-szklanych

1. Wstęp

Etap wymiarowania elementów konstrukcyjnych wymaga uwzględnienia rodzaju materiałów oraz charakteru ich pracy dla oddziałującego obciążenia zewnętrznego (zależność σ - ε wartości naprężeń w funkcji odkształceń) definiowanego stałymi materiałowymi, tj. moduł sprężystości podłużnej (E, moduł Young'a), moduł sprężystości poprzecznej (G, moduł Kirchhoff'a), współczynnik sprężystości objętościowej (K, moduł Helmholtz'a dla materiałów izotropowych), współczynnik Poisson'a (ν , liczba Poisson'a) oraz stałe Lamégo (μ i λ dla materiałów izotropowych).

Popularne w praktyce projektowej jest przyjmowanie z tablic inżynierskich modułów Young'a jako wartości stałych dla danej klasy wytrzymałościowej kompozytu cementowego (podejście będące formą uproszczenia algorytmu wymiarowania i kształtowania elementów nośnych). Stosowana przez projektantów-konstruktorów praktyka zakładania niezmienności wartości modułu sprężystości podłużnej dla danej klasy wytrzymałościowej kompozytu nie jest prawidłowym podejściem, często świadczącym o braku zrozumienia idei charakterystyki materiałowej oraz zasad nauki obejmującej wytrzymałość mechaniczną kompozytów. Identyczną klasę wytrzymałościową materiału, lecz o zmiennej wartości modułu Young'a, można uzyskać dla grupy lub rodziny kompozytów zaprojektowanych z różnej ilości oraz proporcji składników dozowanych w odniesieniu do masy lub objętości wszystkich komponentów receptury.

Wartość modułu Young'a zależna jest od zastosowanych składników w strukturze projektowanego materiału, do których (mających największy wpływ) można zaliczyć rodzaj kruszywa, jego ilość masową lub objętościową frakcji lub grup frakcji, genezę, pochodzenie, kategorię oraz wytrzymałość mechaniczną. Należy podkreślić rolę udziałów procentowych frakcji kruszyw w odniesieniu do ilości spoiwa lub układu spoiw, dodatków wiążących lub wypełniających, jak również masy domieszek i procentowej zawartości porów powietrznych (stopnia napowietrzenia) w strukturze stwardniałego kompozytu cementowego.

Prowadzone badania naukowe w zakresie ustalenia wpływu poszczególnych składników odpadowych, porecyklingowych stanowią szczególnie ważny obszar inżynierii materiałowej oraz technologii produkcji materiałów i wyrobów budowlanych [1-12]. Jak pokazują najnowsze wyniki badań laboratoryjnych nie tylko zastosowane kruszywo, ale również technologia wykonania kompozytów w różnych warunkach termiczno-wilgotnościowych decyduje o wartości współczynnika sprężystości podłużnej [13-16]. Szczególnie istotne staje się prowadzenie badań doświadczalnych dla kompozytów cementowych niewymagających ingerencji człowieka na etapie zagęszczania mieszanki,

¹ waldemarlasic@gmail.com, badacz niezależny.

tj. technologia mieszanek samozagęszczalnych lub prawie samozagęszczalnych (SCC i ASCC) [17-20].

Celem pracy jest określenie wpływu dodatku odpadowych wiórów stalowych o zmiennej geometrii na wartości modułu Young'a dla stwardniałych ekokompozytów cementowo-szklanych oraz dokonanie oceny przydatności odpadów stalowych jako potencjalnego modyfikatora struktury wewnętrznej materiału (ekokompozyt cementowo-szklany jako nowa forma zagospodarowania odpadowych wiórów stalowych z procesów obróbki półproduktów metalowych).

2. Charakterystyka składników ekokompozytu cementowo-szklanego

Główny składnik konstrukcji stosu okruszowego stanowiły grupy frakcji granulatów szkła sodowego oraz mączki szkła sodowego stanowiącej wypełniacz (uszczelniając) struktury ekokompozytu cementowo-szklanego. Frakcje i grupy frakcji szkła sodowego pochodziły z procesów recyklingu odpadowej stłuczki szklanej sodowej, tj. bezbarwne i zielone transparentne szkło postkonsumpcyjne zużytych opakowań produktów spożywczych (butelki i słoiki, kruszone szkło GLS71 oraz GLS72). Odpadową stłuczkę szklaną sodową, przetworzono metodą mechanicznego kruszenia, mielenia i frakcjonowania szkła do wymaganej frakcji lub grupy frakcji. Proces mechanicznego kruszenia poprzedzono oczyszczaniem z etykiet papierowych i syntetycznych. Granulaty szklane poddano zabiegowi trzykrotnego płukania w celu usunięcia pyłu szklanego otaczającego ziarna grupy frakcji 0,063/2,0 mm oraz 2,0/8,0 mm (wzrost przyczepności matrycy cementowej do ziaren granulatu szklanego pozbawionego frakcji pylastej).

Stosy okruszowe zaprojektowano jako w trzech wariantach o uziarnieniu 0/2,0 mm; 0/4,0 mm (granulaty drobne) oraz 0/8,0 mm (granulaty o ciągłym uziarnieniu, mieszanka granulatu drobnego 0/4,0 mm i grubego 4,0/8,0 mm). Grupy frakcji granulatów szklanych sodowych wraz z doziarnieniem (uszczelnieniem) w postaci mączki szkła sodowego stanowiły pełną konstrukcję stosów okruszowych (nie stosowano kruszyw mineralnych naturalnych i kruszyw skalnych łamanych). Granulat drobnej frakcji 0,500/1,0 mm, otrzymaną metodą implozji stłuczki szklanej, stanowiło ścierniwo szklane wykorzystywane w procesach obróbki strumieniowej powierzchni stalowych, betonowych i drewnianych (granulat pozyskany od zakładów produkujących profesjonalne ścierniwa techniczne stosowanych w procesach piaskowania i szkiełkowania). Ziarna szklane posiadały wyraźne krawędzie, powierzchnie ziaren granulatu o odłamach muszlowych bez mikrorys powierzchniowych i mikropęknięć strukturalnych. Analiza składu granulometrycznego granulatu grupy frakcji 0/8,0 mm (wykonana metodą „na sucho”), wykazała procentową zawartość nadziarna (frakcja 8,0/16 mm) w zakresie 1,8-2,3% masy próbki reprezentatywnej (naważka 2000 ± 2 g) oraz zawartość ziaren płaskich w zakresie 1,3-1,7% masy próbki reprezentatywnej (ziarna soczewkowe po kruszeniu denek butelek, pomiar i oznaczenie kształtu ziaren szkła suwmiarką Schultz'a frakcji 8,0/16,0 mm). Wykonany przesiew granulatów szkła sodowego pozwolił na odseparowanie ziaren nieforemnych i nadziaren szklanych frakcji 8,0/16,0 mm. Przyjęto właściwości fizykochemiczne granulatów szkła sodowego: twardość według skali Mohsa równa 6, $\text{SiO}_2 + \text{Na}_2\text{O} \geq 85,0\%$ oraz gęstość objętościowa $\rho_g = 2500 \text{ kg/m}^3$ [21, 22].

Projekt szklanych stosów okruszowych zakładał użycie dodatku mączki szkła sodowego grup frakcji 0/0,200 mm w charakterze uszczelniaacza (doziarnienia) struktury wewnętrznej ekokompozytu cementowo-szklanego. Asortyment szkła sodowego

o uziarnieniu 0,2 mm; 0/4,0 mm i 0/8,0 mm posiadał braki frakcyjne 0/0,063 mm i 0,063/0,125 mm, wynikające z płukania po zakończonych procesach mechanicznego kruszenia i mielenia. Mączkę szkła sodowego pozyskano w wyniku mielenia odpadowego szkła bezbarwnego, transparentnego zielonego i brązowego w postaci nadziaren szklanych, ziaren nieforemnych oraz ziaren płaskich pochodzących z fragmentów denek butelek i słoików. Analizator rozkładu cząstek szklanych pozwolił na określenie zakresów procentowych frakcji szkła dla próbki reprezentatywnej: 0/0,010 mm – 20,0%; 0,010/0,025 mm – 25,0%; 0,025/0,063 mm – 15,0%; 0,063/0,125 mm – 25,0%; 0,125/0,200 mm – 15,0%. Przyjęto parametry fizykochemiczne mączki szkła sodowego: twardość według skali Mohsa równa 6, $\text{Na}_2\text{O} + \text{K}_2\text{O} = 13,0\text{-}15,0\%$; $\text{Al}_2\text{O}_3 = 0,5\text{-}2,0\%$; $\text{SiO}_2 = 70,0\text{-}74,0\%$; $\text{Fe}_2\text{O}_3 \leq 0,1\%$; $\text{CaO} = 7,0\text{-}11,0\%$; $\text{MgO} = 3,0\text{-}5,0\%$; $\text{TiO}_2 \leq 0,1\%$ oraz gęstość objętościowa $\rho_m = 2500 \text{ kg/m}^3$ [23, 24].

Rolę spoiwa hydraulicznego pełnił cement biały specjalny, rodzaju CEM I (cement portlandzki) klasy wytrzymałościowej 52,5 R NA SR5, spoiwo niskoalkaliczne charakteryzujące się wyższą odpornością na działanie siarczanów [25]. Cement specjalny (C) stanowił mieszankę spoiwową wraz z dodatkiem drobno mielonego metakaolinitu (Mk) o udziale procentowym równym 40,0% C + 60,0% Mk (redukcja masy spoiwa cementowego rodzaju CEM I oraz wprowadzenie częściowego substytutu cementu w postaci metakaolinitu). Wytrzymałość rzeczywistą na ściskanie spoiwa cementowego CEM I określono w zakresie 63,3-68,5 MPa po 28 dniach dojrzewania próbek belkowych (wyznaczona dla próbek połówek belek zaprojektowanych na mieszance cementu rodzaju CEM I, wody zarobowej efektywnej i piasku normowego). Przyjęto parametry fizykochemiczne spoiwa cementowego specjalnego: $\text{SO}_3 = 1,8\text{-}2,3\%$; $\text{MgO} = 0,6\%$; $\text{Na}_2\text{O} \leq 0,3\%$; $\text{Cl} \leq 0,04\%$; gęstość objętościowa $\rho_c = 3100 \text{ kg/m}^3$ oraz mialkość równa $400 \text{ m}^2/\text{kg}$ [26]. Zaczyn spoiwowy w układzie cement-metakaolinit wykonano przy użyciu wody zarobowej efektywnej (w_{eff}) o wartości pH = 7,2-7,3 (temperatura wody na etapie mieszania zaczynu spoiwowego w zakresie $+17,0 \pm 2,0^\circ\text{C}$; temperatura powietrza w zakresie $+18,0 \pm 2,0^\circ\text{C}$ oraz wilgotność względna powietrza w zakresie $46,0 \pm 2,0\%$) [27].

Zastosowano częściowy substytut spoiwa cementowego rodzaju CEM I w postaci dodatku metakaolinitu (Mk), składnika o właściwościach pucolanowych aktywowanego wodą zarobową efektywną (dodatek wiążący typu II wdrażany do składów mieszanek betonowych zawierających spoiwa cementowe specjalne lub powszechnego użytku). Parametry fizykochemiczne drobno mielonego metakaolinitu: $\text{Cl} \leq 0,05\%$; $\text{Na}_2\text{O}_{\text{eq}} \leq 1,0\%$; $\text{SiO}_2 = 48,0\text{-}58,0\%$; $\text{P}_2\text{O}_5 \leq 0,3\%$; $\text{MgO} \leq 2,0\%$; $\text{Al}_2\text{O}_3 = 37,0\text{-}45,0\%$; $\text{Fe}_2\text{O}_3 = 0,7\text{-}2,3\%$; $\text{TiO}_2 = 0,8\text{-}2,4\%$; powierzchnia właściwa równa $20000 \text{ cm}^2/\text{g}$ oraz gęstość właściwa $\rho_{\text{Mk}} = 2600 \text{ kg/m}^3$ (warunki środowiskowe użycia dodatku: temperatura powietrza w zakresie $+18,0 \pm 2,0^\circ\text{C}$ oraz wilgotność względna powietrza w zakresie $45,0 \pm 2,0\%$) [28].

Redukcję efektywnej wody zarobowej (w_{eff}) mieszanek cementowo-szklanych uzyskano, stosując domieszkę chemiczną płynną na bazie polikarboksylianów (superplastyfikator silnie upłynniający, znacznie redukujący masę wody zarobowej, otrzymany z przetwórstwa odpadów pochodzenia rolniczego) [29, 30]. Przyjęto gęstość objętościową superplastyfikatora w zakresie $\rho_d = 1,07 \pm 0,02 \text{ g/cm}^3$ i wartości pH w zakresie $5,0 \pm 1,0$. Główny cel użycia domieszki chemicznej wiązał się z możliwością uzyskania projektowanej (założonej wstępnie) klasy konsystencji S2/S3 mieszanki cementowo-szklanej przy zmniejszonych wartościach wskaźników wodnych (redukcja wodożądności całko-

witej szklanego stosu okruczowego oraz poprawa właściwości reologicznych, tj. wzrost urabialności mieszanki, ochrona przed segregacją frakcji szkła sodowego). Domieszkę dozowano w zakresie $0,60 \pm 0,02\%$ (zakres dozowania zalecany: 0,50-5,00%; zakres dozowania optymalny: 1,00-2,50%) masy układu spoiwa cement-metakaolinit (40,0%C + 60,0%Mk). Wartość temperatury dozowania oraz mieszania domieszki ustalono w zakresie $+21,4 \pm 0,1^\circ\text{C}$ (wilgotność względna powietrza w zakresie $48,0 \pm 2,0\%$).

Częściowy substytut granulatu szklanego grubego (frakcja 4,0/8,0 mm) stanowił dodatek wiórów stalowych łukowych luźnych, elementowych i spiralnych (pojedyncze uzwojenie). Odpadowe wióry stalowe pozyskano z zakładów specjalizujących się w obróbce (skrawanie, frezowanie) półproduktów metalowych i z tworzyw sztucznych dla branży samochodowej, maszynowej i lotniczej (technologia CNC wieloosiowego frezowania 3D, 4D i 5D części metalowych). Przyjęto gęstość wiórów stalowych równą 7850 kg/m^3 , kształt wiórów jako łukowy lub spiralny (pojedyncza spirala), długość wiórów w zakresie 5,0-12,0 mm (wióry oczyszczone z chłodziwa stosowanego w procesach skrawania). Dodatek odpadowy dozowano w ilościach 6,0% (113 kg/m^3); 7,0% (132 kg/m^3); 8,0% (150 kg/m^3); 9,0% (169 kg/m^3); 10,0% (188 kg/m^3); 12,0% (226 kg/m^3) oraz 14,0% (263 kg/m^3) względem masy granulatu stosu okruczowego.

3. Metodyka projektowania ekokompozytu cementowo-szklanego

Zastosowano analizę wielokryterialną umożliwiającą dokonanie wyboru metody doboru jakościowego i ilościowego składników mieszanki cementowo-szklanej modyfikowanej odpadowym dodatkiem w postaci wiórów stalowych. Kierowano się algorytmami obliczeniowymi bazującymi na kryteriach wykorzystania minimalnej masy referencyjnego spoiwa cementowego gwarantującego zachowanie wysokiej wytrzymałości oraz trwałości materiału kompozytowego. Uwagę skupiono na projektach składów mieszanek samozagęszczalnych (SCC lub ASCC), opartych na minimalnej ilości zaczynu spoiwowego (metoda Vana i Montgomery'ego) lub metodzie japońskiej [31-35], tj. metoda Okamury, Ozawy i zespołu (dobór ilościowy frakcji stosu okruczowego z uwzględnieniem zjawiska blokowania przepływu mieszanki przez układ gęsto rozmieszczonych prętów zbrojeniowych). Rozważano współczesne metody projektowania mieszanek cementowych drobnokruszywowych wieloskładnikowych, tj. piaskokompozyty oraz kompozyty cementowe o podwyższonym punkcie piaskowym [36-43]. Dokonano przeglądu literatury krajowej (zakres technologii betonów cementowych) oraz artykułów naukowych poruszających zagadnienie innowacyjnego projektowania fibrokompozytów [44-46], kompozytów z uwzględnieniem składników odpadowych [47-56] lub porecyklingowych [57-61]. Przykład stanowią publikacje opisujące kompozyty cementowe średniej wytrzymałości (metoda trzech lub czterech równań) [62-67], wysokiej wytrzymałości (HPC, metoda pojedynczego lub podwójnego otulenia kruszywa). Oddzielną rodzinę analizowanych materiałów wysokiej wytrzymałości stanowiły ekokompozyty cementowe wykorzystujące założenia koncepcji współczynnika aktywności pucolanowej lub hydraulicznej (koncepcja współczynnika „k” [68-72], ECPC lub EPCC, zastosowanie odpadowych spoiw przemysłowych jako częściowych substytutów cementu). Szczególną uwagę skupiono na metodach projektowania kompozytów specjalnych o wysokiej trwałości i szczelnym i gęstym rozmieszczeniu (model ścisłego upakowania CMP, klasy dominujące) wypełniaczy pylastych, drobnych frakcji kruszyw i dodatków wiążących w stosie okruczowym, tj. ultrawysokowartościowe betony (UHPC)

oraz kompozyty modyfikowane proszkami reaktywnymi (RPC, Eco-RPC, Green-RPC, SHCC) [73-77].

Składy receptur mieszanek cementowo-szklanych zaprojektowano metodą czterech równań z modyfikacją w postaci nowej autorskiej koncepcji opartej na równoległym zastosowaniu dwóch współczynników aktywności pucolanowej lub hydraulicznej (koncepcja współczynników „k2” i „k3”). Zmodyfikowana metoda umożliwiła dokonanie doboru jakościowego i ilościowego składu mieszanek referencyjnych i modyfikowanych dodatkiem odpadowych wiórów stalowych. Założenia metody projektowej uwzględniały rozwiązania proekologiczne, tj. redukcja masy referencyjnej spoiwa cementowego (wprowadzenie dwóch odpadowych spoiw przemysłowych do składu mieszanki), zastosowanie mikrowłókien lub makrowłókien w postaci odpadowego kordu (stalowego lub tekstylnego) celem poprawy wytrzymałości na rozciąganie przy zginaniu i rozłupywaniu (zasada *zero waste*, ograniczenie stosowania konwencjonalnego zbrojenia rozproszonego). Ideą nowej metody projektowania jest uzyskanie składów mieszanek cementowo-szklanych generujących na etapie produkcji zmniejszoną emisję dwutlenku węgla (redukcja śladu węglowego w cyklu życia wyrobu). Główny cel metody stanowi częściowe lub pełne odstępianie od stosowania surowców w składach eko-kompozytów, tj. kruszyw mineralnych naturalnych lub łamanych skalnych oraz wysoko klinkierowych spoiw hydraulicznych rodzaju CEM I. Drugorzędny cel autorskiej metody projektowania eko-kompozytów cementowo-szklanych ukierunkowany został na wdrożenie innowacyjnych rozwiązań nowej rodziny eko-kompozytów wysokiej wytrzymałości wykonanych w technologii nisko cementowej lub bez cementowej przy zachowaniu wysokiej trwałości materiału (ekofibrokompozyty oraz lekkie ekostyrokompozyty cementowo-szklane z równoważnymi właściwościami użytkowymi).

Główne założenie projektowe eko-kompozytu cementowo-szklanego stanowiło uzyskanie materiału wysokiej wytrzymałości badanej po 28 dniach dojrzewania próbek cylindrycznych ($f_{ck,cyl} \geq 60,0$ MPa). Rozważano możliwość uzyskania konstrukcji stosów okrucowych wykonanych wyłącznie z granulatów szklanych sodowych z uszczelnieniem w postaci wyselekcjonowanych frakcji pochodzących z mączki szkła sodowego (użycie szklanego kruszywa łamanego porecyklingowego i/lub odpadowego, projektowana krzywa uziarnienia szklanego stosu okrucowego). Założono klasy ekspozycji (warunki oddziaływania środowiskowego) przy ustalaniu granicznych wartości składników mieszanki cementowo-szklanej, tj. XC3, XD2, XF2, XA1 oraz XM1. Przyjęto minimalną masę spoiwa cementowego specjalnego SR ($C > 300$ kg/m³), maksymalny współczynnik wodno-spoiwowy ($w/s < 0,55$), minimalną klasę wytrzymałości na ściskanie (klasa stwardniałego kompozytu $>C30/37$) oraz minimalną zawartość powietrza w mieszance cementowo-szklanej ($p \geq 5,5\%$) [78, 79].

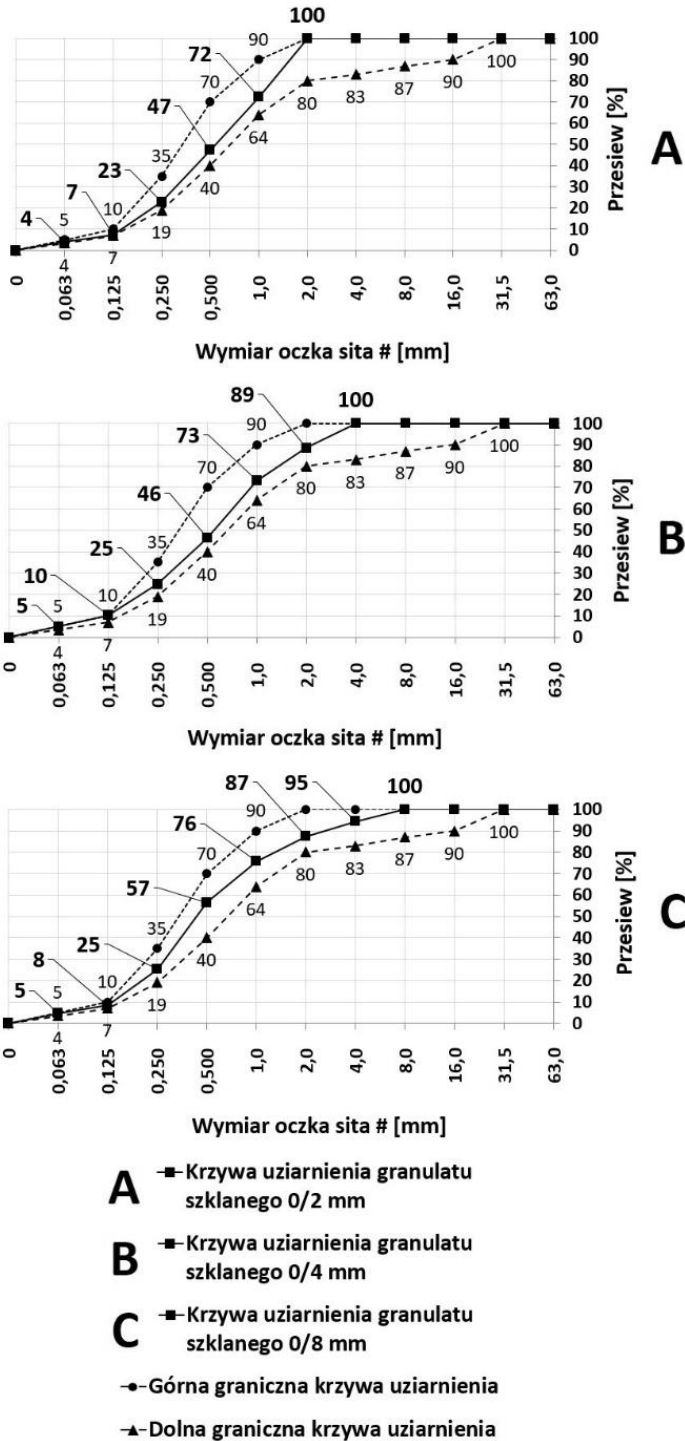
Projektowane krzywe uziarnienia szklanych stosów okrucowych odniesiono do granicznych (dolnej i górnej) krzywych uziarnienia jak dla piaskokompozytów i betonów piaskowych o podwyższonym punkcie piaskowym. Wyznaczono wskaźniki wodne (wodożądność frakcji) dla łamanego granulatu szklanego drobnego (w_{gd}) i grubego (w_{gg}), spoiwa cementowego (w_c) według tablic O. Stern'a w zależności od założonej klasy konsystencji (S3, średnia klasa ciekłości) [80-83]. Dokonano korekty wartości wskaźników wodnych (w_d i w_g) dla frakcji granulatów szkła sodowego wynikającą z różnicy gęstości objętościowej i rodzaju powierzchni ziaren (granulat łamany) względem kruszywa mineralnego naturalnego. Wykorzystano nomogram z krzywymi Waltz'a

umożliwiający określenie wartości wskaźnika wodno-cementowego (w/c) lub wodno-spoiwowego (w/s) dla założonej klasy wytrzymałości na ściskanie stwardniałego ekoekompozytu. Obliczono masy składników mieszanki na podstawie równań metody projektowej. Pierwsze równanie (warunek wytrzymałościowy) umożliwiło wyznaczenie średniej wartości wytrzymałości na ściskanie badanej po 28 dniach dojrzewania próbek (f_{cm}). Równanie drugie (warunek szczelności) pozwoliło wyznaczyć sumę objętości absolutnej składników mieszanki (uwzględnienie objętości porów powietrznych, $p \geq 5,5\%$). Warunek konsystencji (urabialności mieszanki) stanowił równanie trzecie metody projektowej, gwarantował obliczenie ilości efektywnej wody zarobowej (W_{eff}) dla klasy konsystencji S2/S3. Objętość zaprawy w ekoekompozycie cementowo-szklanym wyznaczono na podstawie wzoru czwartego, określono wskaźniki spulchnienia kruszywa szklanego drobnego i grubego (zaprawę stanowiła mieszanka zaczynu spoiwowego i drobnego granulatu szklanego grup frakcji 0/2,0 mm lub 0/4,0 mm). Metodą krzywych granicznych zaprojektowano szklane krzywe uziarnienia 0/2,0 mm; 0/4,0 mm oraz 0/8,0 mm. Przyjęto wartości punktów charakterystycznych, tj. Pp – punkt pyłowy, PPP – punkt pyłowo-piaskowy oraz PP – punkt piaskowy w zakresach: (uziarnienie 0/2,0 mm) Pp = 7-10%; PPP = 42-55%; PP = 100%; (uziarnienie 0/4,0 mm) Pp = 8-10%; PPP = 40-60%; PP = 85-95% oraz (uziarnienie 0/8,0 mm) Pp = 7-10%; PPP = 50-65%; PP = 82-93%. Założono maksymalny wymiar ziarna granulatu szklanego $D_{max} = 8,0$ mm.

Zaprojektowane krzywe uziarnienia stosów okruchowych o uziarnieniu 0/2,0 mm; 0/4,0 mm oraz 0/8,0 mm przedstawiono na rysunku 1.

4. Metodyka badawcza

Serie próbek ekoekompozytu cementowo-szklanego, podlegające badaniu wytrzymałości charakterystycznej na ściskanie ($f_{ck,cyl}$) oraz modułu sprężystości podłużnej ($E_{c,s}$), pielęgnowano w warunkach normowych przez okres 28 dni od momentu formowania (temperatura wody w procesie dojrzewania próbek w zakresie $18,0 \pm 3,0^\circ\text{C}$; temperatura powietrza w zakresie $18,0^\circ\text{C} \pm 2,0^\circ\text{C}$; wilgotność względna powietrza w zakresie $47,0 \pm 2,0\%$) [84]. Badania prowadzono na stwardniałych próbkach cylindrycznych o wysokości 300,0 mm (wysokość rzeczywista w zakresie 294,0-296,0 mm), średnicy podstawy górnej oraz dolnej równej 150,0 mm (średnica rzeczywista w zakresie 149,0-150,0 mm; zmiany wynikające ze skurczu chemicznego zaczynu w układzie cement-metakaolinit oraz skurczu autogenicznego). Odchyłki wysokości próbek cylindrycznych podyktowane zostały względami technologicznymi, tj. proces cięcia mechanicznego górnej powierzchni próbki zapewniający równość i płaskość powierzchni podlegających obciążeniu powierzchniowemu (zapewnienie warunku równomiernego rozkładu obciążenia powierzchniowego, brak potrzeby wykonania warstwy wyrównującej w postaci zaprawy z cementu glinowego i piasku drobnego) [85]. Dokonano oceny wizualnej stwardniałych próbek ekoekompozytu, sprawdzono podstawowe właściwości fizyczne próbek cementowo-szklanych (masa, gęstość w stanie pełnego nasycenia wodą). Próbki niespełniające wymagań w zakresie kształtu i wymiarów odrzucono ze zbioru reprezentatywnego.



Rysunek 1. Krzywe uziarnienia szklanych stosów okruszowych eko kompozytów cementowo-szklanych. Wykresy: A – krzywa uziarnienia 0/2,0 mm; B – krzywa uziarnienia 0/4,0 mm; C – krzywa uziarnienia 0/8,0 mm [opracowanie własne]

Badanie wytrzymałości charakterystycznej na ściskanie ($f_{ck,cyl}$) oraz wyznaczenie średniej wytrzymałości na ściskanie ($f_{cm,cyl}$) przeprowadzono na seriach stwardniałych próbek cylindrycznych po 28 dniach dojrzewania w warunkach normowych (próbki cylindryczne formowane) [86]. Zastosowano maszynę wytrzymałościową typu Controls 3000 kN (klasa I) z napędem hydraulicznym [87]. Sprawdzone prawidłowość działania czujników przemieszczeń liniowych tłoka ściskającego oraz wartości przemieszczeń punktów kontrolnych próbki testowej (próbka kontrolna niepodlegająca badaniu). Wybrano zakres pomiarowy siły obciążającej równy 150-3000 kN. Osuszone z wody powierzchniowej próbki cylindryczne zamocowano w pionowej przestrzeni roboczej maszyny wytrzymałościowej. Dokonano centrowania osi próbek względem pionowej osi układu płyt ściskających (odniesienie położenia próbki cylindrycznej względem koncentrycznie rozmieszczonych kołowych znaczników płyty dolnej i górnej, redukcja mimośrodków wynikających z błędnego usytuowania badanej próbki). Prawidłowe zamocowanie próbek cylindrycznych stanowi wymóg konieczny gwarantujący wystąpienie wyłącznie efektu ściskania osiowego bez pojawienia się negatywnych naprężeń rozciągających. Ustalono prędkość naprężeń w funkcji czasu w zakresie $0,6 \pm 0,2$ MPa/s. Zastosowano obciążenie wstępne generujące wartość naprężeń o wartości w zakresie $\sigma = 5,0 \pm 1,0$ MPa. Statyczna próba ściskania trwała do momentu zniszczenia próbek cylindrycznych, tj. powstanie odprysków powierzchniowych oraz rys niszczących (spadek wartości siły maksymalnej F_{max} poniżej $20,0 \pm 2,0\%$). Porównano formę zniszczenia próbki cylindrycznej do form określonych zapisami normowymi, obliczono wartość wytrzymałości charakterystycznej na ściskanie dla próbek cylindrycznych ($f_{ck,cyl}$) oraz średnią wartość wytrzymałości charakterystycznej na ściskanie ($f_{cm,cyl}$) dla serii próbek cylindrycznych. Oceniono formę zniszczenia próbek na prawidłową, wyznaczono wartości naprężeń dolnych $\sigma_b = (0,10-0,20) \cdot f_{cm,cyl}$; naprężeń górnych $\sigma_a = f_{cm,cyl}/3$ oraz naprężeń ściskających wstępnych równych $\sigma_p = 0,5$ MPa dla normowej metody badania „A” lub z zakresu $0,5 \text{ MPa} \leq \sigma_p \leq \sigma_b$ dla normowej metody badania „B”).

Podstawę wyznaczenia wartości siecznego ustabilizowanego modułu Young'a ($E_{c,s}$) przy ściskaniu stanowiło badanie określające wartości przemieszczeń liniowych dla próbek cylindrycznych 150 x 300 mm. Rzeczywiste wymiary próbek cylindrycznych spełniały warunek ilorazu wysokości do średnicy próbki w zakresie 2-4 oraz warunku minimalnej średnicy próbki względem maksymalnego wymiaru ziarna granulatu $D_{cyl} > 3,5 \cdot D_{max}$ ($D_{max} = 8,0$ mm) [88]. Zastosowano dwa samonośne ekstensometry o bazach pomiarowych równych 150,0 mm i 2,0 mm, tj. średnicowy ekstensometr ramkowy prostokątny typu Epsilon 3975 klasy 0,2 (błąd względny $q_{Le} = \pm 0,2\%$) oraz uśredniający osiowy ekstensometr typu Epsilon 3542RA klasy 0,5 (błąd względny $q_{Le} = \pm 0,5\%$) [89, 90]. Zastosowany w badaniach modułu sprężystości podłużnej układ dwóch ekstensometrów typu Epsilon 3542RA i 3975, użyto w konfiguracji 3542RA-150M-600M-ST (zakres maksymalnego dopuszczalnego pionowego przemieszczenia liniowego $\pm 6,0$ mm) oraz w konfiguracji 3975-200M-ST (zakres maksymalnych dopuszczalnych poziomych przemieszczeń do $+2,0$ mm) [91, 92]. Pomiar wartości przemieszczeń liniowych w kierunku pionowym oraz poziomym względem działającej siły dokonano odpowiednio ekstensometrem średnicowym pionowym i ramkowym poziomym. Dokonano regulacji średnicy objęcia ekstensometru typu Epsilon 3542RA względem średnicy każdej próbki cylindrycznej podlegającej badaniu. Określono wartości przemieszczeń liniowych, obliczono wartości odkształceń liniowych podłużnych

dla siecznego modułu Young'a i odkształceń liniowych poprzecznych dla współczynnika Poisson'a.

Wartości odkształceń liniowych wyznaczono poprzez iloraz mierzonych wartości przemieszczeń liniowych podłużnych oraz poprzecznych do długości bazy pomiarowej ekstensometrów ($\varepsilon = \Delta L/L_0$, $\Delta L = L_1 - L_0$, $L_0 = 150,0$ mm dla przemieszczeń liniowych pionowych, $L_0 = 2,0$ mm dla przemieszczeń liniowych poziomych). Prawdliwość pracy ekstensometru osiowego i średnicowego (wpływającej na uzyskane wartości odczytów) zapewniono przez utrzymanie stabilnych warunków środowiskowych, tj. monitorowana i rejestrowana wartość temperatury i wilgotności względnej powietrza w zakresie $+18,0^\circ\text{C} \pm 2,0^\circ\text{C}$ (dokładność odczytu $\pm 0,5^\circ\text{C}$ dla termohigrometru, odczyty mierzonych wartości przemieszczeń liniowych zawierały się w zakresach temperatury prawidłowej pracy ekstensometrów od $-40,0^\circ\text{C}$ do $+100,0^\circ\text{C}$). Badanie poprzedzono operacją kalibracji ekstensometrów za pomocą kalibratora typu Epsilon 3590AT o rozdzielczości 0,001 mm [93], wyposażonego w zestaw adapterów, kolumn rozszerzających i wsporników montażowych (weryfikacja liniowego charakteru pracy urządzenia dla analizowanego zakresu pomiarowego). Stożkowe styki nośne ekstensometrów zamocowano na szlifowanej i odtłuszczonej powierzchni poboczniczy próbek cylindrycznych (redukcja negatywnego wpływu drgań maszyny wywołującej efektu poślizgu i niekontrolowanego przesuwu ekstensometrów wzdłuż tworzących walec prosty). Wartości modułu Young'a przeprowadzono metodą „A” oraz „B”. Dokonano obliczeń wartości siecznych ustabilizowanych modułów sprężystości podłużnej $E_{c,s}$ według wzorów danej metody (1) i (2).

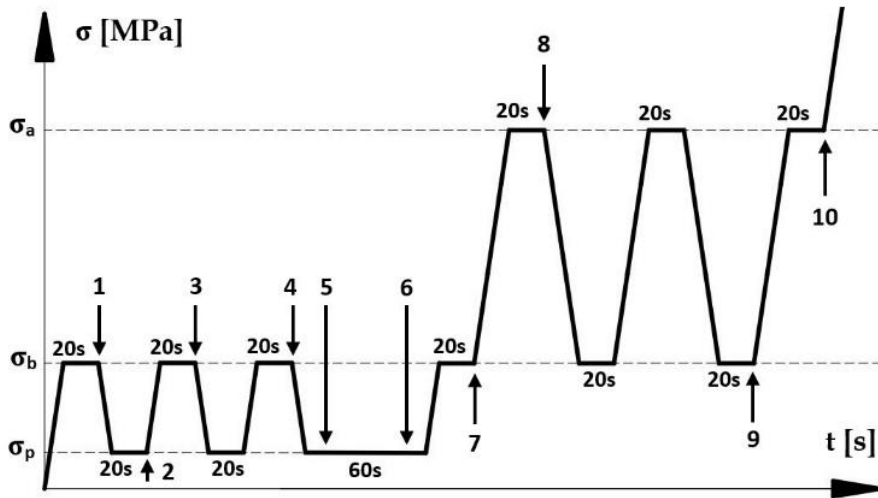
$$(1) E_{c,s} = \frac{(\sigma_{a,3}^m - \sigma_{b,2}^m)}{(\varepsilon_{a,3} - \varepsilon_{b,2})} \text{ (metoda „A”)}$$

$$(2) E_{c,s} = \frac{(\sigma_{a,3}^m - \sigma_{p,2}^m)}{(\varepsilon_{a,3} - \varepsilon_{p,2})} \text{ (metoda „B”)}$$

Opis symboliki użytej we wzorach: $\sigma_{a,3}^m$ – zmierzona wartość naprężenia odpowiadająca nominalnemu górnemu naprężeniu σ_a [MPa] zarejestrowana po zakończonym cyklu trzecim obciążania (przed rozpoczęciem etapu obciążania do momentu zniszczenia próbki); $\sigma_{b,2}^m$ – zmierzona wartość naprężenia odpowiadająca nominalnemu dolnemu naprężeniu σ_b [MPa] zarejestrowana po zakończonym cyklu drugim obciążania próbki; $\varepsilon_{a,3}$ – średnie odkształcenie przy górnym naprężeniu w cyklu trzeciego obciążania [wartość bezwymiarowa]; $\varepsilon_{b,2}$ – średnie odkształcenie przy dolnym naprężeniu w cyklu drugiego obciążania [wartość bezwymiarowa]; $\sigma_{p,2}^m$ – nominalna wartość naprężenia wstępnego zarejestrowana po drugim cyklu obciążania, wartość z zakresu od 0,5 MPa do wartości $\leq \sigma_b$ [MPa] zarejestrowana po zakończeniu drugiego cyklu obciążania próbki; $\varepsilon_{p,2}$ – średnie odkształcenie przy wstępnym naprężeniu w cyklu drugiego obciążania [wartość bezwymiarowa].

Metoda „A” pozwoliła określić początkowy moduł Young'a ($E_{c,0}$) oraz umożliwiła wyznaczenie wartości siecznego ustabilizowanego modułu Young'a ($E_{c,s}$) uwzględniającego wartość naprężeń dolnych σ_b i uzyskanych wartości odkształceń liniowych ε_b . Metoda „B” umożliwiła wyznaczenie siecznego ustabilizowanego modułu sprężystości podłużnej $E_{c,s}$ (moduł $E_{c,s}$ uwzględniający wartość naprężeń wstępnych σ_p i uzyskanych wartości odkształceń liniowych ε_p).

Algorytm postępowania wraz z opisem punktów charakterystycznych metody badawczej „A”, określający wartości siecznego ustabilizowanego modułu sprężystości podłużnej $E_{c,s}$ przedstawiono na rysunku 2.



$$\sigma_a = f_{cm,cyl} / 3$$

$$\sigma_b = (0,10 - 0,20) \cdot f_{cm,cyl}$$

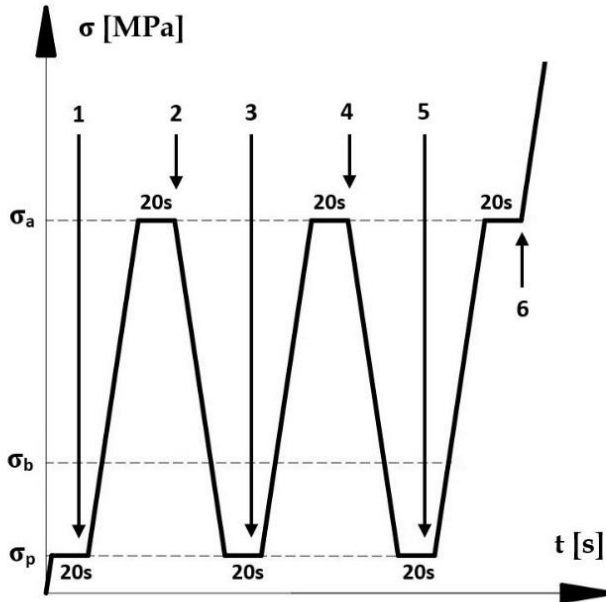
$$\sigma_p = 0,5 \text{ MPa}$$

$$E_{c,s} = (\sigma_{a,3}^m - \sigma_{b,2}^m) / (\varepsilon_{a,3} - \varepsilon_{b,2})$$

1.	Zapis wartości σ_b^m .
2.	Wyzerowanie ekstensometrów.
3.	Wyznaczenie wartości $\varepsilon_{b,2i}$.
4.	Wyznaczenie wartości $\varepsilon_{b,3i}$.
5.	Sprawdzenie warunku (sprawdzenie pierwsze): $\varepsilon_{b,3i} - \varepsilon_{b,2i} \leq 10\%$.
6.	Sprawdzenie warunku (sprawdzenie drugie): $\varepsilon_{b,3i} - \varepsilon_{b,2i} \leq 20\%$.
7.	Wyznaczenie wartości $\varepsilon_{b,0i}$, wyznaczenie średniej wartości $\varepsilon_{b,0}$, zapis wartości $\sigma_{b,0}^m$.
8.	Wyznaczenie wartości $\varepsilon_{a,1i}$, wyznaczenie średniej wartości $\varepsilon_{a,1}$, zapis wartości $\sigma_{a,1}^m$.
9.	Wyznaczenie wartości $\varepsilon_{b,2i}$, wyznaczenie średniej wartości $\varepsilon_{b,2}$, zapis wartości $\sigma_{b,2}^m$.
10.	Wyznaczenie wartości $\varepsilon_{a,3i}$, wyznaczenie średniej wartości $\varepsilon_{a,3}$, zapis wartości $\sigma_{a,3}^m$.

Rysunek 2. Schemat przebiegu badania metodą „A” wraz z opisem punktów charakterystycznych [opracowanie własne]

Algorytm postępowania wraz z opisem punktów charakterystycznych metody badawczej „B”, określający wartości siecznego ustabilizowanego modułu sprężystości podłużnej $E_{c,s}$ przedstawiono na rysunku 3.



$$\sigma_a = f_{cm,cyl} / 3$$

$$\sigma_b = (0,10 - 0,20) \cdot f_{cm,cyl}$$

$$0,5 \text{ MPa} \leq \sigma_p \leq \sigma_b$$

$$E_{C,S} = (\sigma_{a,3}^m - \sigma_{p,2}^m) / (\varepsilon_{a,3} - \varepsilon_{p,2})$$

1.	Wyzerowanie ekstensometrów.
2.	Wyznaczenie wartości $\varepsilon_{a,1i}$, wyznaczenie średniej wartości $\varepsilon_{a,1}$.
3.	Wyznaczenie wartości $\varepsilon_{p,1i}$, wyznaczenie średniej wartości $\varepsilon_{p,1}$.
4.	Wyznaczenie wartości $\varepsilon_{a,2i}$, wyznaczenie średniej wartości $\varepsilon_{a,2}$, sprawdzenie warunku (sprawdzenie pierwsze): $\varepsilon_{a,2i} - \varepsilon_{a,1} \leq 20\%$.
5.	Wyznaczenie wartości $\varepsilon_{p,2i}$, wyznaczenie średniej wartości $\varepsilon_{p,2}$, zapis wartości $\sigma_{p,2}^m$.
6.	Wyznaczenie wartości $\varepsilon_{a,3i}$, wyznaczenie średniej wartości $\varepsilon_{a,3}$, zapis wartości $\sigma_{a,3}^m$, sprawdzenie warunku (sprawdzenie drugie): $\varepsilon_{a,3} - \varepsilon_{a,2} \leq 10\%$.

Rysunek 3. Schemat przebiegu badania metodą „B” wraz z opisem punktów charakterystycznych [opracowanie własne]

5. Analiza wyników

Serie próbek podlegające badaniu wytrzymałości charakterystycznej na ściskanie ($f_{ck,cyl}$; wyznaczenie wartości wytrzymałości $f_{cm,cyl}$) oraz sieciowego modułu sprężystości podłużnej ($E_{c,s}$) oznaczono nazwami przedstawionymi w tabeli 1.

Tabela 1. Zastosowane nazewnictwo próbek ekokompozytu cementowo-szklanego referencyjnego i podlegającego modyfikacji materiałowej w postaci dodatku odpadowych stalowych wiórów łukowych i elementowych, dozowanych względem referencyjnej masy granulatu szkła sodowego stosu okruczowego (% m. g.)

Lp.	Nazwa serii próbek	Zawartość procentowa odpadowych wiórów stalowych łukowych luźnych i elementowych dla danej serii próbek ekokompozytów cementowo-szklanych [% m. g.]
1.	R	0% (0 kg/m ³)
2.	P1	6,0% (113 kg/m ³)
3.	P2	7,0% (132 kg/m ³)
4.	P3	8,0% (150 kg/m ³)
5.	P4	9,0% (169 kg/m ³)
6.	P5	10,0% (188 kg/m ³)
7.	P6	12,0% (226 kg/m ³)
8.	P7	14,0% (263 kg/m ³)

Źródło: Opracowanie własne.

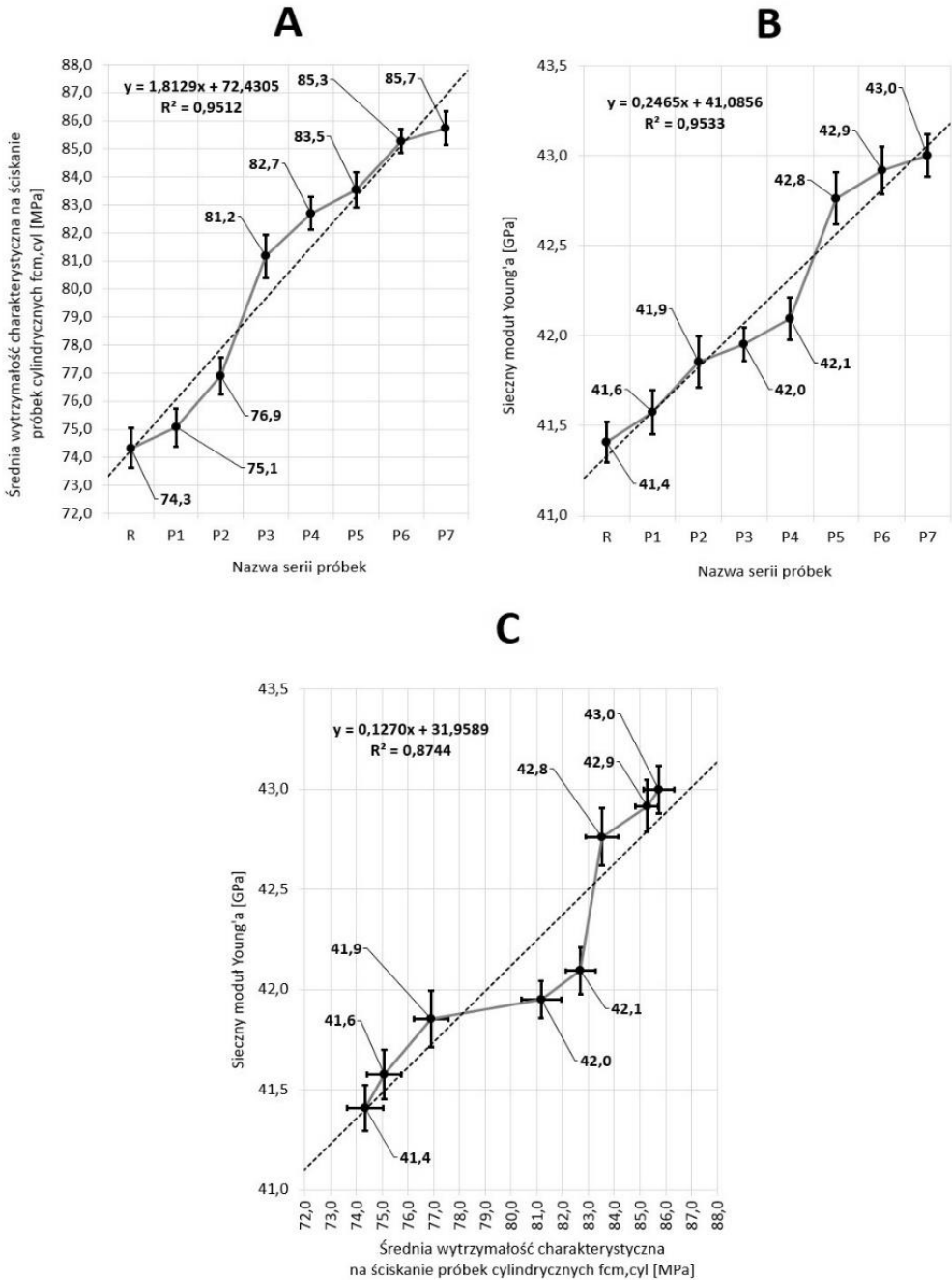
Badanie wytrzymałości charakterystycznej na ściskanie ($f_{ck,cyl}$) przeprowadzono na próbkach cylindrycznych po 28 dniach dojrzewania ekokompozytu cementowo-szklanego. Określono wartości $f_{ck,cyl}$ dla 3 próbek dla każdej serii (populacji), wyznaczono średnią wartość wytrzymałości charakterystycznej na ściskanie z trzech próbek cylindrycznych ($f_{cm,cyl}$) dla serii próbek referencyjnych oraz modyfikowanych dodatkiem odpadowych wiórów stalowych. Wykonano dwa zestawy próbek złożonych z ośmiu serii, tj. zestaw „A” oraz zestaw „B”. Wartości wytrzymałości $f_{cm,cyl}$ (określonej dla serii R, P1, P2, P3, P4, P5, P6 i P7) wykorzystano w metodzie „A” i w metodzie „B” do wyznaczania wartości siecznego modułu Young’a (całkowita liczba próbek równa 48 dla dwóch zestawów serii).

Serie referencyjne próbek cylindrycznych uzyskały średnią wartość wytrzymałości $f_{cm,cyl}$ równą 74,3 MPa (seria referencyjna „A”) oraz 75,0 MPa (seria referencyjna „B”). Próbkki referencyjne charakteryzowały się najmniejszą uzyskaną wytrzymałością $f_{cm,cyl}$ ze wszystkich badanych serii ekokompozytów cementowo-szklanych. Wartości siecznego modułu sprężystości podłużnej ($E_{c,s}$) próbek serii referencyjnej (R) ekokompozytu cementowo-szklanego wyniosły odpowiednio 41,4 GPa (metoda badawcza „A”) oraz 40,6 GPa (metoda badawcza „B”).

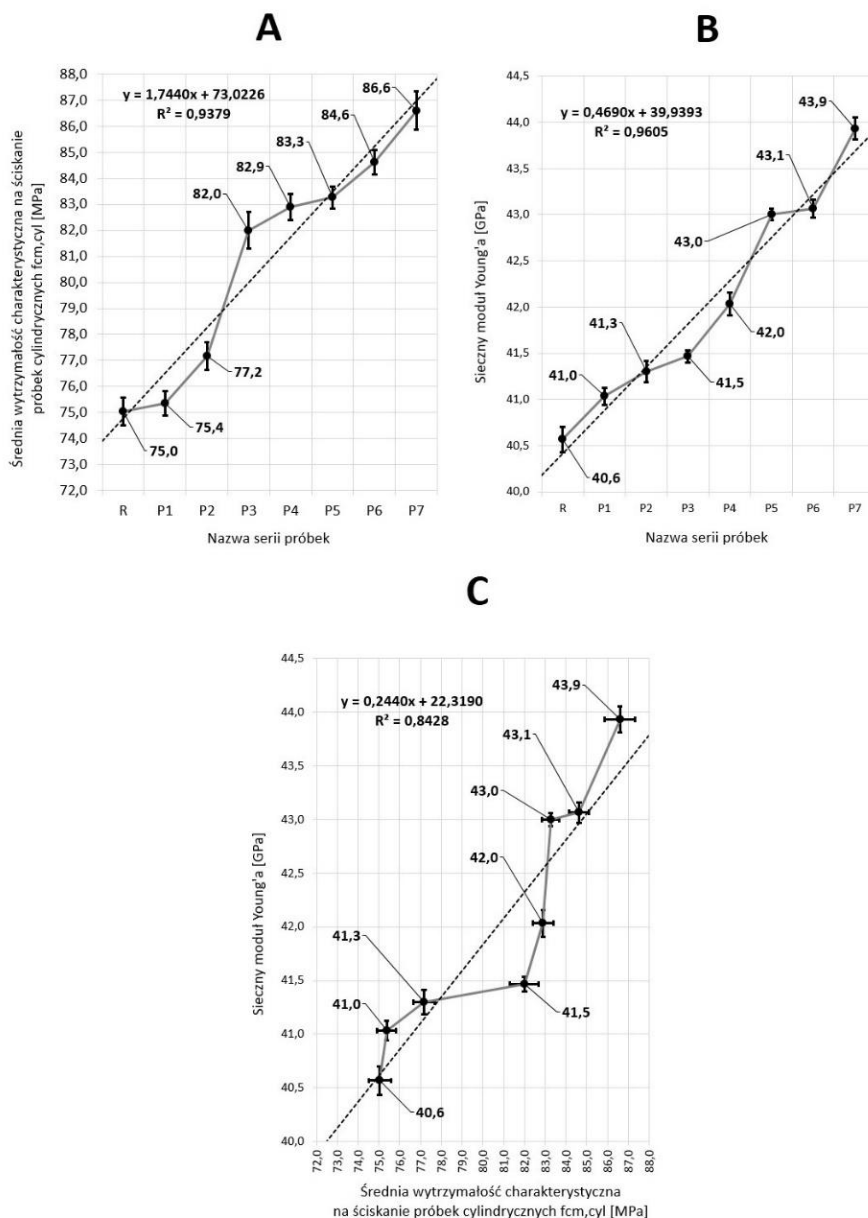
Wyniki badań wytrzymałości charakterystycznej na ściskanie ($f_{ck,cyl}$ wraz z wyznaczeniem wytrzymałości $f_{cm,cyl}$) wyznaczone po 28 dniach dojrzewania ekokompozytu cementowo-szklanego oraz wartości modułu Young’a (określone metodą „A”) przedstawiono na rysunku 4.

Największą wartość wytrzymałości $f_{cm,cyl}$ określono dla próbek serii P7 dla zestawu próbek „A” i „B” (wytrzymałość $f_{cm,cyl}$ równa 85,7 MPa oraz 86,6 MPa). Zaobserwowano wzrost wartości wytrzymałości $f_{cm,cyl}$ w zakresie 15,3-15,5% względem próbek serii referencyjnej.

Wyniki badań wytrzymałości charakterystycznej na ściskanie ($f_{ck,cyl}$ wraz z wyznaczeniem wytrzymałości $f_{cm,cyl}$) wyznaczone po 28 dniach dojrzewania ekokompozytu cementowo-szklanego oraz wartości modułu Young’a (określone metodą „B”) przedstawiono na rysunku 5.



Rysunek 4. Wyniki badań wytrzymałości mechanicznej oraz wartości modułu Young'a (metoda „A”) po 28 dniach dojrzewania eko kompozytu cementowo-szklanego: A – średnia wytrzymałość charakterystyczna na ściskanie ($f_{cm,cyl}$) próbek cylindrycznych, B – wartości modułu Young'a wyznaczone metodą „A”, C – wykres zależności wartości siecznego modułu sprężystości podłużnej (E_{cs}) w funkcji średniej wartości wytrzymałości charakterystycznej na ściskanie ($f_{cm,cyl}$) próbek cylindrycznych [opracowanie własne]



Rysunek 5. Wyniki badań wytrzymałości mechanicznej oraz wartości modułu Young'a (metoda „B”) po 28 dniach dojrzewania eko kompozytu cementowo-szklanego: A – średnia wytrzymałość charakterystyczna na ściskanie ($f_{ck,cyl}$) próbek cylindrycznych, B – wartości modułu Young'a wyznaczone metodą „A”, C – wykres zależności wartości siecznego modułu sprężystości podłużnej (E_{cs}) w funkcji średniej wartości wytrzymałości charakterystycznej na ściskanie ($f_{cm,cyl}$) próbek cylindrycznych [opracowanie własne]

Największą wartość modułu Young'a zaobserwowano dla próbek serii P7 równe odpowiednio 43,0 GPa (metoda badacza „A”) oraz 43,9 GPa (metoda badacza „B”).

Podstawowe wartości statystyk opisowych mierzonych wielkości dla próbek eko kompozytu cementowo-szklanego przedstawiono na rysunku 6.

Nazwa serii próbek	Rodzaj badania	Statystyki opisowe																										
		N-ważnych	Średnia	Ufności (-95,000%)	Ufności (95,000%)	Średnia Wiersora (5,0000%)	Geometr. (Średnia)	Harmon. (Średnia)	Mediana	Minimum	Maksimum	Dolny (Kwartyl.)	Górny (Kwartyl.)	Percentyl (10,00000)	Percentyl (90,00000)	Rozstęp	Kwartyl. (Rozstęp)	Wariancja	Odchylenie standardowe	P. ufności odch. std. (-95,000%)	P. ufności odch. std. (+95,000%)	Współczynnik zmienności	Standard. (Błąd)	Skośność	Bł. std. (Skośność)			
R	Statyczna próba ściskania próbek cylindrycznych - wytrzymałość charakterystyczna na ściskanie fck.cyl.[MPa] - metoda "A"	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
P1		3	74,3	69,7	78,9	74,3	74,3	74,3	74,3	74,3	73,8	72,8	76,4	72,8	76,4	72,8	76,4	3,60	3,60	3,45	1,86	0,97	11,68	2,50	1,07	1,19	1,22	
P2		3	75,1	70,9	79,1	75,1	75,1	75,1	75,1	75,1	75,6	73,2	76,4	73,2	76,4	73,2	76,4	3,20	3,20	2,77	1,67	0,87	10,47	2,22	0,96	-1,29	1,22	
P3		3	76,9	73,0	80,8	76,9	76,9	76,9	76,9	76,9	76,3	75,7	78,7	75,7	78,7	75,7	78,7	3,00	3,00	2,52	1,59	0,83	9,98	2,06	0,92	1,46	1,22	
P4		3	81,2	75,8	86,6	81,2	81,2	81,1	81,1	81,1	80,5	79,4	83,6	79,4	83,6	79,4	83,6	4,20	4,20	4,74	2,18	1,13	13,69	2,68	1,26	1,25	1,22	
P5		3	82,7	79,6	85,8	82,7	82,7	82,7	82,7	82,7	83,2	81,3	83,6	81,3	83,6	81,3	83,6	2,30	2,30	2,30	1,51	1,23	6,44	1,72	1,49	0,71	-1,53	1,22
P6		3	83,5	79,8	87,3	83,5	83,5	83,5	83,5	83,5	84,5	82,1	85,1	82,1	85,1	82,1	85,1	3,00	3,00	2,26	1,50	0,78	9,45	1,80	0,87	1,40	1,22	
P7	3	85,3	83,6	86,9	85,3	85,3	85,3	85,3	85,3	85,6	84,5	85,7	84,5	85,7	84,5	85,7	1,20	1,20	0,44	0,67	0,35	4,18	0,78	0,38	0,40	1,22		
R	Moduł sprężystości podłużnej - moduł Younga dla próbek cylindrycznych Ecs.[GPa] - metoda "A"	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
P1		3	41,4	39,3	43,5	41,4	41,4	41,4	41,4	41,4	41,1	40,7	42,3	40,7	42,3	40,7	42,3	1,60	1,60	0,69	0,83	0,43	5,23	2,01	0,48	1,30	1,22	
P2		3	41,9	39,3	44,4	41,9	41,9	41,8	41,8	42,0	40,7	42,8	40,7	42,8	40,7	42,8	40,7	42,8	2,05	2,05	1,07	1,03	0,54	6,49	2,47	0,60	-0,69	1,22
P3		3	42,0	40,2	43,7	42,0	42,1	42,1	42,1	42,0	41,3	43,0	41,3	43,0	41,3	43,0	41,3	43,0	1,72	1,72	0,75	0,86	0,45	5,43	2,05	0,50	0,48	1,22
P4		3	42,8	40,1	45,4	42,8	42,8	42,8	42,7	42,2	42,1	44,0	42,1	44,0	42,1	44,0	42,1	44,0	1,91	1,91	1,12	1,06	0,55	6,66	2,48	0,61	1,68	1,22
P5		3	42,9	40,5	45,3	42,9	42,9	42,9	42,9	42,9	43,5	41,8	43,5	41,8	43,5	41,8	43,5	1,68	1,68	0,92	0,96	0,50	6,03	2,24	0,56	-1,73	1,22	
P6		3	43,0	40,8	45,2	43,0	43,0	43,0	43,0	42,6	42,4	44,0	42,4	44,0	42,4	44,0	42,4	44,0	1,61	1,61	0,76	0,87	0,45	5,49	2,03	0,50	1,61	1,22
R	Statyczna próba ściskania próbek cylindrycznych - wytrzymałość charakterystyczna na ściskanie fck.cyl.[MPa] - metoda "B"	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
P1		3	75,0	72,4	77,6	75,0	75,0	75,0	75,0	74,7	74,2	76,2	74,2	76,2	74,2	76,2	74,2	76,2	2,00	2,00	1,08	1,04	0,54	6,54	1,39	0,60	1,29	1,22
P2		3	75,4	73,3	77,4	75,4	75,4	75,4	75,4	75,0	74,8	76,3	74,8	76,3	74,8	76,3	74,8	76,3	1,50	1,50	0,66	0,81	0,42	5,12	1,08	0,47	1,62	1,22
P3		3	77,2	74,7	79,7	77,2	77,2	77,2	77,1	76,2	78,2	76,2	78,2	76,2	78,2	76,2	78,2	2,00	2,00	1,00	1,00	0,52	6,30	1,30	0,58	0,30	1,22	
P4		3	82,0	77,5	86,5	82,0	82,0	82,0	81,5	80,5	84,0	80,5	84,0	80,5	84,0	80,5	84,0	3,50	3,50	3,25	1,80	0,94	11,33	2,20	1,04	1,15	1,22	
P5		3	82,9	80,5	85,3	82,9	82,9	82,9	82,5	82,2	84,0	82,2	84,0	82,2	84,0	82,2	84,0	1,80	1,80	0,93	0,96	0,50	6,06	1,16	0,56	1,55	1,22	
P6		3	83,3	81,7	84,9	83,3	83,3	83,3	83,0	82,8	84,0	82,8	84,0	82,8	84,0	82,8	84,0	1,20	1,20	0,41	0,64	0,33	4,04	0,77	0,37	1,55	1,22	
P7	3	84,6	82,6	86,7	84,6	84,6	84,6	85,0	83,7	85,2	83,7	85,2	83,7	85,2	83,7	85,2	1,50	1,50	0,66	0,81	0,42	5,12	1,08	0,47	-1,62	1,22		
R	Moduł sprężystości podłużnej - moduł Younga dla próbek cylindrycznych Ecs.[GPa] - metoda "B"	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
P1		3	40,6	38,1	43,1	40,6	40,6	40,6	40,6	40,5	39,6	41,6	39,6	41,6	39,6	41,6	39,6	41,6	2,00	2,00	1,00	1,00	0,52	6,30	2,47	0,68	0,30	1,22
P2		3	41,0	39,3	42,7	41,0	41,0	41,0	40,8	40,5	41,8	40,5	41,8	40,5	41,8	40,5	41,8	1,30	1,30	0,46	0,68	0,35	4,28	1,66	0,39	1,36	1,22	
P3		3	41,3	40,2	43,4	41,3	41,3	41,3	41,3	41,4	40,5	42,2	40,5	42,2	40,5	42,2	40,5	42,2	1,00	1,00	0,73	0,85	0,44	5,37	2,01	0,29	0,59	1,22
P4		3	42,0	39,7	44,3	42,0	42,0	42,0	42,3	41,6	42,8	41,6	42,8	41,6	42,8	41,6	42,8	1,80	1,80	0,66	0,93	0,48	5,84	2,21	0,54	-1,19	1,22	
P5		3	43,0	41,9	44,1	43,0	43,0	43,0	43,0	42,9	42,6	43,5	42,6	43,5	42,6	43,5	42,6	43,5	0,90	0,90	0,21	0,46	0,24	2,88	1,07	0,26	0,94	1,22
P6		3	43,1	41,3	44,8	43,1	43,1	43,1	43,1	43,0	42,4	43,8	42,4	43,8	42,4	43,8	42,4	43,8	1,40	1,40	0,49	0,70	0,37	4,41	1,63	0,41	0,42	1,22
P7	3	43,9	41,7	46,2	43,9	43,9	43,9	43,9	44,0	43,0	44,8	43,0	44,8	43,0	44,8	43,0	44,8	1,80	1,80	0,81	0,90	0,47	5,67	2,05	0,52	-0,33	1,22	

Rysunek 6. Wartości statystyk opisowych: średnia wytrzymałość charakterystyczna na ściskanie próbek cylindrycznych ($f_{cm,cyl}$); wartości modułu Young'a dla metody badawczej „A” i „B” [opracowanie własne]

6. Dyskusja

Dokonano porównania uzyskanych wyników badań własnych względem danych zawartych w zagranicznych artykułach naukowych. Wyselekcjonowane publikacje dotyczyły problematyki modyfikacji materiałowej mikrostruktury kompozytów cementowych wraz z określeniem wpływu danej modyfikacji na wartości stałych materiałowych, tj. modułu Young'a oraz współczynnika Poisson'a. Odniesiono się do wyników badań fibrobetonów cementowych zbrojonych fazą włóknistą (nanowłókna węglowe CNF) oraz względem drobnokruszywowych zapraw cementowych o zmiennych proporcjach i konfiguracjach składników wiążących (spoiwa cementowe i wapienne hydrauliczne) oraz kruszyw (piaski rzeczne). Cytowane artykuły naukowe o zbliżonej tematyce wybrano ze względu na zauważalne braki publikacji naukowych dotyczących badań laboratoryjnych modułu sprężystości podłużnej dla konstrukcyjnych kompozytów cementowych modyfikowanych wiórami stalowymi z procesów obróbki w technologii CNC półproduktów stalowych.

Określeniem wpływu międzyfazowej strefy przejściowej dla matrycy spoiwowej i powierzchni ziaren kruszywa na wartość modułu sprężystości podłużnej oraz współczynnika Poisson'a zajmowali się Zhu X. wraz z zespołem badawczym w publikacji naukowej [94]. Obszar rozważań naukowców koncentrował się na badaniach laboratoryjnych w zakresie wytrzymałości mechanicznej strefy przejściowej spoiwo-kruszywo (strefa ITZ) z wykorzystaniem metody elementów skończonych dla modelu międzyfazowego. Zhu X. wraz z zespołem zaproponowali dwa modele opisujące strefę przejściową, tj. model międzyfazowy (*Interphase Model*) oraz model sferycznej warstwy wewnętrznej sprężystej z wariantami obciążeń promienistych i obwodowych (*Spring Layer Interface Model*). Model strefy ITZ opisany dodatkową warstwą międzyfazową, ograniczony matrycą z zaprawy cementowej i „wtrąceniem” w postaci kruszywa grubego. Przeprowadzono symulację metodą elementów skończonych (FEM) modelu strefy przejściowej ITZ, uwzględniająca wpływ grubości strefy ITZ na wytrzymałość mechaniczną kompozytu cementowego [94]. Skład betonu stanowiło spoiwo cementowe rodzaju CEM I, piasek rzeczny, kruszywo grube łamane wapienne oraz domieszka nanowłókien węglowych (CNF) dozowanych w ilościach 0,05% i 0,1% zaprojektowanej masy cementu ($w/c = 0,51$; stosunek piaskowo-cementowy równy 2,46). Dodatek modyfikujący skład referencyjny kompozytów cementowych stanowiły nanowłókna węglowe średnicy 100 nm, długości w zakresie 50-200 nm będące nową formą zbrojenia rozproszonego podnoszącego wytrzymałość mechaniczną strefy przejściowej ITZ, tj. redukcja mikrorys w matrycy spoiwowej oraz ograniczenie negatywnych nanoporów i mikroporów powietrznych istotnych w teorii mechaniki pękania materiałów anizotropowych. Wyniki badań cylindrycznych próbek zaprawy oraz betonu (według normy ASTM) wykazały wzrost wytrzymałości charakterystycznej na ściskanie $f_{ck,cyl}$ w zakresie statycznego oddziaływania obciążenia zewnętrznego. Zaobserwowano wzrost wartości modułu Young'a o 33,7% (dla $f_{cm,cyl} = 43,1$ MPa, próbki 100 x 200 mm) dla zaprawy cementowej oraz dla betonu cementowego o 39,6% względem serii referencyjnej (dla $f_{cm,cyl} = 46,2$ MPa, dozowanie nanowłókien węglowych CNF w ilości 0,1% zaprojektowanej masy spoiwa hydraulicznego) [94]. Domieszkowanie nanomateriałem generowało powstawanie aglomeracji, stanowiących lokalne imperfekcje mikrostruktury wewnętrznej (nieciągłości strukturalne) w betonie cementowym. Autor rozdziału uzyskał trend wzrostowy wartości średniej wytrzymałości charakterystycznej na ściskanie $f_{cm,cyl}$

w zakresie 1,1-15,3% oraz 0,5-15,5% odpowiednio dla metod pomiarowych „A” i „B” (rysunek 4 i 5). Według badań własnych, dodatek odpadowych wiórów stalowych dla wszystkich modyfikacji wpłynął na wzrost wartości siecznego modułu Young’a ($E_{c,s}$) w zakresie 41,4-43,0 GPa dla metody pomiarowej „A” oraz w zakresie 40,6-43,9 GPa według metody pomiarowej „B” (rys. 4 i 5).

Problematykę analizy odkształcalności próbek zapraw murarskich w badaniu modułu Young’a metodą statyczną i dynamiczną poruszył Marques A.I. wraz z zespołem badawczym [95]. Badania naukowców koncentrowały się na analizie, kontroli i redukcji wartości odkształceń wpływających na rozwój rys niszczących oraz na ich dalszą propagację w próbkach belkowych stwardniałego kompozytu cementowego. Zaproponowano nową, eksperymentalną metodę oznaczania wartości modułu sprężystości podłużnej zaadaptowanej względem klasycznej metody pomiarowej dla próbek betonowych [95]. Wykorzystano metodę ultradźwiękową (DME-US, *Direct Ultrasonic Methodology*) oraz metodę dynamicznej częstotliwości rezonansowej (*Resonance Frequency dynamic methodology*) dla dziesięciu wariantów receptur zapraw murarskich (oznaczenie dynamicznego modułu sprężystości podłużnej oraz prędkości rozchodzenia się podłużnych lub poprzecznych fal ultradźwiękowych w kompozycie cementowym). Próbki belkowe 40 x 40 x 160 mm zaprojektowano na spoiwie cementowym oraz na wapnie hydraulicznym, modyfikacje materiałowe określone zostały zmiennymi proporcjami spoiw i kruszyw (uzyskano kompozyty zawierające piasek rzeczny o gęstości objętościowej w zakresie 1971-2070 kg/m³, stosunek spoiwa cementowego lub wapiennego hydraulicznego do piasku określono na 1 : 2; 1 : 3 oraz 1 : 4) [95]. Badaniu podlegały próbki po 14, 21, 28, 90 i po 365 dniach dojrzewania, uzyskano wartości wytrzymałości charakterystycznej na ściskanie równe odpowiednio 23,3 MPa; 28,2 MPa; 25,0 MPa; 28,2 MPa oraz 27,6 MPa. Największe wartości statycznego modułu Young’a wykazały próbki zapraw murarskich dla proporcji spoiwa do kruszywa równej 1 : 2 (24-26 GPa dla cementu badanych po 14 i 28 dniu dojrzewania; 5,8-6,5 GPa dla wapna hydraulicznego dla 28 i 365 dnia dojrzewania próbek belkowych). Badania przeprowadzone przez Marques A.I. wraz z zespołem badawczym wykazały istotną zależność dla proporcji rozpatrywanego spoiwa, a masą kruszywa naturalnego na zmianę wartości modułu Young’a. Zwiększone dozowanie spoiwa cementowego lub wapiennego hydraulicznego kosztem redukcji masy kruszywa naturalnego wpływało na wzrost wartości modułu Young’a w odniesieniu do zapraw o zwiększonej zawartości kruszywa (proporcja 1 : 3 spoiwa do kruszywa naturalnego) [95]. Odnosząc się do badań własnych prowadzonych metodą „A” oraz „B”, stwierdzono trend wzrostowy wartości wytrzymałości charakterystycznej na ściskanie próbek cylindrycznych, jak i wzrost wartości siecznego modułu Young’a w funkcji procentowego dozowania odpadowych wiórów stalowych (rys. 4 i 5).

7. Podsumowanie

Dodatek odpadowych wiórów stalowych łukowych luźnych, elementowych i spiralnych, dozowanych w ilościach 6,0%; 7,0%; 8,0%; 9,0%; 10,0%; 12,0% oraz 14,0% masy układu granulatu szkła sodowego, wpłynął na zmiany w zakresie:

- właściwości reologicznych mieszanek cementowo-szklanych:
 - obniżenie urabialności mieszanki dla dozowania wiórami stalowymi w ilościach 10,0%; 12,0% i 14,0% (przyrost czasu zagęszczania mechanicznego mieszanki w zakresie 25-45 sekund w odniesieniu do próbki mieszanki referencyjnej),

- obniżenie klasy konsystencji (ciekłości mieszanki) z S3 na S2 (dozowanie w zakresie 6,0-9,0%) i z S3 na S1 (dozowanie w zakresie 10,0-14,0%),
- efekt osiadania wiórów w dolnych warstwach próbki w wyniku dłuższego czasu zagęszczania mechanicznego przy obniżonej urabialności (wypieranie granulatu szklanego do górnych warstw mieszanki cementowo-szklanej),
- powstawanie aglomeracji wiórów stalowych w dolnych warstwach zagęszczanej mieszanki (efekt rozwarstwienia próbki dla dozowania odpadów w zakresie 10,0-16,0%; niejednorodny rozkład wiórów w objętości próbki),
- wzrost wskaźnika wodno-spoiwowego (w/s) w układzie cement-metakaolinit w zakresie 0,38-0,42;
- zwiększone dozowanie domieszki chemicznej (superplastyfikatora na bazie polikarboksylanów), wzrost procentowej zawartości domieszki upłynniającej w zakresie 0,65-0,68% (dla temperatury dozowania i mieszania równej $+21,4^{\circ}\text{C} \pm 0,1^{\circ}\text{C}$) masy układu spoiwa cement-metakaolinit,
- ograniczenie migracji zaczynu spoiwowego w układzie cement-metakaolinit do górnych warstw mieszanki dla dozowania wiórami stalowymi w zakresie 6,0-9,0% (brak wpływów i zacieków zaczynu spoiwowego cement-metakaolinit na poboczniczy próbek cylindrycznych),
- mikrostruktury po 28 dniach dojrzewania próbek cylindrycznych:
 - zwiększenie procentowej zawartości negatywnych makroporów powietrznych o średnicy $>300\ \mu\text{m}$ ($>A300$, zakres makroporów: A300-2700 o zmiennym kształcie), zwiększenie negatywnego napowietrzenia w ilościach 2,7%; 2,8% i 3,1% dla dozowania wiórami w ilościach 10,0%; 12,0% oraz 16,0%,
 - nierównomierny rozkład negatywnych makroporów powietrznych A300-2700 w trzech rozpatrywanych przekrojach stwardniałej próbki (wskaźnik rozmieszczenia makroporów „L” w zakresie 0,23-0,27; analizowane przekroje mikrostruktury ekokompozytu na wysokościach $50 \pm 3\ \text{mm}$; $150 \pm 3\ \text{mm}$ oraz $250 \pm 3\ \text{mm}$ określone w odniesieniu od dolnej podstawy próbki),
- wytrzymałości mechanicznej po 28 dniach dojrzewania próbek cylindrycznych:
 - metoda „A”: wzrost średniej wytrzymałości charakterystycznej na ściskanie ($f_{\text{cm,cyl}}$) w zakresie dozowania 6,0-14,0% dodatku wiórów stalowych o 1,1-15,3% względem próbek cylindrycznych serii referencyjnej,
 - metoda „B”: wzrost średniej wytrzymałości charakterystycznej na ściskanie ($f_{\text{cm,cyl}}$) w zakresie dozowania 6,0-14,0% dodatku wiórów stalowych o 0,5-15,5% względem próbek cylindrycznych serii referencyjnej,
- siecznego modułu Young'a, tj. stałej materiałowej ($E_{\text{c,s}}$) określonej po 28 dniach dojrzewania próbek cylindrycznych:
 - metoda „A”: wzrost średniej wartości siecznego modułu sprężystości podłużnej ($E_{\text{c,s}}$) w zakresie dozowania 6,0-14,0% dodatku wiórów stalowych o 0,5-3,9% względem próbek cylindrycznych serii referencyjnej,
 - metoda „B”: wzrost średniej wartości siecznego modułu sprężystości podłużnej ($E_{\text{c,s}}$) w zakresie dozowania 6,0-14,0% dodatku wiórów stalowych o 1,0-8,1% względem próbek cylindrycznych serii referencyjnej.

Literatura

1. Lin H., Takasu K., Suyama H., Koyamada H., Liu S., *A study on properties, static and dynamic elastic modulus of recycled concrete under the influence of modified fly ash*, Construction and Building Materials, 347, 2022, s. 128585.
2. Fuzail Hashmi A., Shariq M., Baqi A., *An investigation into age-dependent strength, elastic modulus and deflection of low calcium fly ash concrete for sustainable construction*, Construction and Building Materials, 283, 2021, s. 122772.
3. Saridemir M., *Effect of silica fume and ground pumice on compressive strength and modulus of elasticity of high strength concrete*, Construction and Building Materials, 49, 2013, s. 484-489.
4. Shariq M., Prasad J., Abbas H., *Effect of GGBFS on age dependent static modulus of elasticity of concrete*, Construction and Building Materials, 41, 2013, s. 411-418.
5. Alengaram U.J., Mahmud H., Jumaat M.Z., *Enhancement and prediction of modulus of elasticity of palm kernel shell concrete*, Materials and Design, 32, 2011, s. 2143-2148.
6. Wongpa J., Kiattikomol K., Jaturapitakkul C., Chindaprasirt P., *Compressive strength, modulus of elasticity, and water permeability of inorganic polymer concrete*, Materials and Design, 31, 2010, s. 4748-4754.
7. Yoshitake I., Zhang W., Mimura Y., Saito T., *Uniaxial tensile strength and tensile Young's modulus of fly-ash concrete at early age*, Construction and Building Materials, 40, 2013, s. 514-521.
8. Bellum R.R., Muniraj K., Madduru S.R.C., *Investigation on modulus of elasticity of fly ash-ground granulated blast furnace slag blended geopolymer concrete*, Materials Today: Proceedings, 27, 2020, s. 718-723.
9. Kalla J., Karri S., Sathi K.V., *Experimental analysis on modulus of elasticity of slag based concrete*, Materials Today: Proceedings, 37, 2021, s. 2114-2120.
10. Han T., Siddique A., Khayat K., Huang J., Kumar A., *An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete*, Construction and Building Materials, 244, 2020, s. 118271.
11. Shao J., Zhu H., Zhao B., Haruna S.I., Xue G., Jiang W., Wu K., Yang J., *Combined effect of recycled tire rubber and carbon nanotubes on the mechanical properties and microstructure of concrete*, Construction and Building Materials, 322, 2022, s. 126493.
12. Yan P., Wu J., Lin D., Liu X., *Uniaxial compressive stress-strain relationship of mixed recycled aggregate concrete*, Construction and Building Materials, 350, 2022, s. 128663.
13. Maruyama I., Sasano H., Nishioka Y., Igarashi G., *Strength and Young's modulus change in concrete due to long-term drying and heating up to 90°C*, Cement and Concrete Research, 66, 2014, s. 48-63.
14. Liu B.D., Lv W.J., Li L., Li P.F., *Effect of moisture content on static compressive elasticity modulus of concrete*, Construction and Building Materials, 69, 2014, s. 133-142.
15. Zhang G., Li C., Wei H., Wang M., Yang Z., Gu Y., *Influence of Humidity on the Elastic Modulus and Axis Compressive Strength of Concrete in a Water Environment*, Materials, 2020, 13, s. 5696.
16. Algourdin N., Pliya P., Beaucour A.L., Noumowe A., di Coste D., *Effect of fine and coarse recycled aggregates on high-temperature behaviour and residual properties of concrete*, Construction and Building Materials, 341, 2022, s. 127847.
17. Parra C., Valcuende M., Gomez F., *Splitting tensile strength and modulus of elasticity of self-compacting concrete*, Construction and Building Materials, 25, 2011, s. 201-207.
18. Vilanova A., Fernandez-Gomez J., Landsberger G.A., *Evaluation of the mechanical properties of self compacting concrete using current estimating models Estimating the modulus of elasticity, tensile strength, and modulus of rupture of self compacting concrete*, Construction and Building Materials, 25, 2011, s. 3417-3426.

19. Craeye B., Van Itterbeek P., Desnerck P., Boel V., De Schutter G., *Modulus of elasticity and tensile strength of self-compacting concrete: Survey of experimental data and structural design codes*, Cement & Concrete Composites, 54, 2014, s. 53-61.
20. Turatsinze A., Garros M., *On the modulus of elasticity and strain capacity of Self-Compacting Concrete incorporating rubber aggregates*, Resources, Conservation and Recycling, 52, 2008, s. 1209-1215.
21. Strona internetowa: <https://pph-rewa.pl/> [data dostępu: 05.09.2022].
22. Strona internetowa: <https://blastgrit.pl/> [data dostępu: 05.09.2022].
23. Strona internetowa: <https://artglas-recykling.pl/> [data dostępu: 05.09.2022].
24. Strona internetowa: <https://artglas-recykling.pl/dokumenty/> [data dostępu: 05.09.2022].
25. PN-B-19707:2013-10, Cement – Cement specjalny – Skład, wymagania i kryteria zgodności, Polski Komitet Normalizacyjny.
26. <https://www.aalborgwhite.com/products-solutions/products/all-products> [data dostępu: 05.09.2022].
27. PN-EN 1008:2004, *Woda zarobowa do betonu – Specyfikacja pobierania próbek, badanie i ocena przydatności wody zarobowej do betonu, w tym wody odzyskanej z procesów produkcji betonu*, Polski Komitet Normalizacyjny.
28. Strona internetowa: <https://astra-polska.com/produkt/dodatki-mineralne/pucolanowe/astramk-40/> [data dostępu: 05.09.2022].
29. Strona internetowa: <https://fra.sika.com/fr/construction/adjuvants-additifs-betons/adjuvants/superplastifiants/sika-viscocrete-850vegetal.html> [data dostępu: 05.09.2022].
30. PN-EN 934-2+A1:2012, *Domieszki do betonu, zaprawy i zaczynu – Część 2: Domieszki do betonu – Definicje, wymagania, zgodność, oznakowanie i etykietowanie*, Polski Komitet Normalizacyjny.
31. Okamura H., Ouchi M., *Self-Compacting Concrete*, Journal of Advanced Concrete Technology, 1(1), 5-15, April 2003.
32. Okamura H., Ozawa K., *Self-Compacting High Performance Concrete*, Structural Engineering International, 6(4), 1996.
33. Okamura H., Ozawa K., *Mix-design for self-compacting concrete*, Concrete Library of JSCE, 25, 1995, s. 107-120.
34. Zuo W., Wei L., Du Z., Li P., Wu Y., She W., *Design of robust Self-Compacting concrete based on analysis of component properties*, Construction and Building Materials, 301, 2021, s. 124063.
35. Li J., Tan D., Zhang X., Wan C., Xue G., *Mixture design method of self-compacting lightweight aggregate concrete based on rheological property and strength of mortar*, Journal of Building Engineering, 43, 2021, s. 102660.
36. Łasica W., Małek M., *Zastosowanie sztuczki szklanej w kompozytach na spoiwie cementowym oraz w betonach*, [w:] *Ochrona środowiska i gospodarka odpadami – współczesne rozwiązania i perspektywy na przyszłość*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 167-187, ISBN 978-83-67104-29-6.
37. Zaitri R., Bederina M., Bouziani T., Makhloufi Z., Hadjoudja M., *Development of high performances concrete based on the addition of grinded dune sand and limestone rock using the mixture design modelling approach*, Construction and Building Materials, 60, 2014, s. 8-16.
38. Łasica W., Małek M., *Wpływ dodatku pyłu szklanego odpadowego na wytrzymałość mechaniczną i właściwości cieplne kompozytów cementowo-szklanych*, [w:] *Ochrona środowiska i gospodarka odpadami – współczesne rozwiązania i perspektywy na przyszłość*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 188-210, ISBN 978-83-67104-29-6.

39. Małek M., Jackowski M., Łasica W., Dydek K., Boczkowska A., *An Experimental Study of Possible Post-War Ferronickel Slag Waste Disposal in Szklary (Lower Silesian, Poland) as Partial Aggregate Substitute in Concrete: Characterization of Physical, Mechanical, and Thermal Properties*, Materials, 14(10), 2021, s. 2552.
40. Łasica W., Małek M., *Wpływ dodatku mączki szklanej oraz odpadowego kordu tekstylnego na wytrzymałość mechaniczną i właściwości cieplne kompozytów cementowo-szklanych*, [w:] *Ochrona środowiska i gospodarka odpadami – współczesne rozwiązania i perspektywy na przyszłość*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 211-240, ISBN 978-83-67104-29-6.
41. Łasica W., Małek M., *Wpływ domieszki biopolimerowej na bazie skrobi na właściwości mieszanek oraz stwardniałych kompozytów cementowo-szklanych*, [w:] *Ochrona środowiska i gospodarka odpadami – współczesne rozwiązania i perspektywy na przyszłość*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 241-266, ISBN 978-83-67104-29-6.
42. Małek M., Grzelak K., Łasica W., Jackowski M., Kluczyński J., Szachogłuchowicz I., Torzewski J., Łuszczek J., *Cement-glass composite bricks (CGCB) with interior 3D printed PET-G scaffolding*, Journal of Building Engineering, March 2022, s. 104429.
43. Spengler A., *Technologie sandreicher Betone*, Lehrstuhl für Baustoffkunde und Werkstoffprüfung der Technischen Universität München, 2006.
44. Małek M., Jackowski M., Łasica W., Kadela M., *Influence of polypropylene, glass and steel fiber on the thermal properties of concrete*, Materials, 14, 2021, s. 1888.
45. Małek M., Jackowski M., Łasica W., Kadela M., Wachowski M., *Mechanical and material properties of mortar reinforced with glass fiber: an experimental study*, Materials, 14, 2021, s. 698.
46. Małek M., Łasica W., Kadela M., Kluczyński J., Dudek D., *Physical and mechanical properties of polypropylene fibre-reinforced cement-glass composite*, Materials, 14, 2021, s. 637.
47. Łasica W., Małek M., *Kompozyt cementowo-szklany jako sposób wykorzystania szkła odpadowego*, [w:] *Współczesne problemy ochrony środowiska – wybrane aspekty*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2021, s. 60-79, ISBN 978-83-66489-86-8.
48. Łasica W., Małek M., *Kompozyt cementowo-szklany modyfikowany metalowymi materiałami odpadowymi*, [w:] *Wybrane zagadnienia z zakresu nanotechnologii, inżynierii materiałowej oraz termodynamiki*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2021, s. 142-158, ISBN 978-83-66489-95-0.
49. Almeshal I., Tayeha B.A., Alyousef R., Alabduljabbar H., Mohamed A.M., *Eco-friendly concrete containing recycled plastic as partial replacement for sand*, Journal of Materials Research and Technology, 9(3), 2020, s. 4631-4643.
50. Łasica W., Małek M., *Wpływ dodatku odpadowych granulowanych łupin orzecha oraz kolb kukurydzy na właściwości mechaniczne kompozytów cementowych i betonów*, [w:] *Nauki ścisłe i przyrodnicze – przegląd wybranych zagadnień*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2021, s. 158-177, ISBN 978-83-67104-13-5.
51. Małek M., Łasica W., Jackowski M., Kadela M., *Effect of waste glass addition as a replacement for fine aggregate on properties of mortar*, Materials, 13, 2020, s. 3189.
52. Małek M., Kadela M., Terpiłowski M., Szewczyk T., Łasica W., Muzolf P., *Effect of Metal Lathe Waste Addition on the Mechanical and Thermal Properties of Concrete*, Materials, 14, 2021, s. 2760.
53. Życiński W., Małek M., Jackowski M., Łasica W., Owczarek M., *Study of the workability and mechanical properties of concrete with added ground corncobs*, Materiali in tehnologije/Materials and technology, 54(4), 2020, s. 479-483.
54. Bhat J.A., *Mechanical behaviour of self compacting concrete: Effect of wood ash and coal ash as partial cement replacement*, Materials Today: Proceedings, 42, 2021, s. 1470-1476.

55. Larsen O., Naruts V., Aleksandrova O., *Self-compacting concrete with recycled aggregates*, Materials Today: Proceedings, 19, 2019, s. 2023-2026.
56. Poongodi K., Murthi P., Gobinath R., *Evaluation of ductility index enhancement level of banana fibre reinforced lightweight self-compacting concrete beam*, Materials Today: Proceedings, 39, 2021, s. 131-136.
57. Małek M., Jackowski M., Łasica W., Kadela M., *Characteristics of Recycled Polypropylene Fibers as an Addition to Concrete Fabrication Based on Portland Cement*, Materials, 13, 2020, s. 1827.
58. Łasica W., Małek M., Szcześniak Z., Owczarek M., *Characterization of recycled glass-cement composite: mechanical strength*, Materiali in tehnologije/Materials and technology, 54(4), 2020, s. 473-477.
59. Guo P., Meng W., Nassif H., Gou H., Bao Y., *New perspectives on recycling waste glass in manufacturing concrete for sustainable civil infrastructure*, Construction and Building Materials, 257, 2020, s. 119579.
60. Jackowski M., Małek M., Życiński W., Łasica W., Owczarek M., *Characterization of new recycled polymer shots addition for the mechanical strength of concrete*, Materiali in tehnologije/Materials and technology, 54, 2020, 3, s. 355-358.
61. Małek M., Życiński W., Jackowski M., Łasica W., Wachowski M., *Effect of recycled fibers addition on mechanical properties of concrete*, Conference: 12th International Conference on Intelligent Technologies in Logistics and Mechatronics Systems, ITELMS 2018, April 2018, Panevezys, Lithuania.
62. Małek M., Jackowski M., Życiński W., Łasica W., Wachowski M., *Characterization on new filler addition for concrete fabrication based on Portland cement*, Conference: 12th International Conference on Intelligent Technologies in Logistics and Mechatronics Systems, ITELMS 2018, April 2018, Panevezys, Lithuania.
63. Małek M., Jackowski M., Życiński W., Łasica W., Wachowski M., *Effect of micrometakaolin addition on mechanical properties of concrete*, Conference: 12th International Conference on Intelligent Technologies in Logistics and Mechatronics Systems, ITELMS 2018, April 2018, Panevezys, Lithuania.
64. Małek M., Jackowski M., Życiński W., Łasica W., Wachowski M., *Mechanical strength characterization of new manufactured concrete with zeolite addition*, Conference: 12th International Conference on Intelligent Technologies in Logistics and Mechatronics Systems, ITELMS 2018, April 2018, Panevezys, Lithuania.
65. Małek M., Życiński W., Jackowski M., Łasica W., Wachowski M., *Effect of new filler addition on the mechanical properties of concrete*, Conference: 12th International Conference on Intelligent Technologies in Logistics and Mechatronics Systems, ITELMS 2018, April 2018, Panevezys, Lithuania.
66. Małek M., Życiński W., Jackowski M., Łasica W., Wachowski M., *Influence of new admixture addition on concrete mechanical properties*, Conference: 12th International Conference on Intelligent Technologies in Logistics and Mechatronics Systems, ITELMS 2018, April 2018, Panevezys, Lithuania.
67. Małek M., Jackowski M., Życiński W., Łasica W., Owczarek M., *Influence of silicone carbide additions on the mechanical properties of concrete*, Materiali in tehnologije/Materials and technology, 54, 2020, 5, s. 595-599.
68. Łasica W., Małek M., *Wpływ dodatku odpadowych mikrowłókien bambusa zwyczajnego na wytrzymałość mechaniczną oraz parametry cieplne kompozytów cementowo-szklanych*, [w:] *Najnowsze doniesienia z inżynierii materiałowej*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 108-135, ISBN 978-83-67104-58-6.
69. Łasica W., Małek M., *Wpływ domieszki odpowietrzającej na właściwości mieszanek oraz stwardniałych kompozytów cementowo-szklanych*, [w:] *Najnowsze doniesienia z inżynierii materiałowej*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 136-157, ISBN 978-83-67104-58-6.

70. Łasica W., Małek M., *Wpływ dodatku odpadowego pyłu polifluorku winylidenu (PVDF) na wytrzymałość mechaniczną oraz parametry cieplne eko kompozytów cementowo-szklanych*, [w:] *Najnowsze doniesienia z inżynierii materiałowej*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 158-183, ISBN 978-83-67104-58-6.
71. Łasica W., Małek M., *Wpływ dodatku odpadowego pyłu polimetakrylanu metylu (PMMA) na wytrzymałość mechaniczną oraz parametry cieplne eko kompozytów cementowo-szklanych*, [w:] *Najnowsze doniesienia z inżynierii materiałowej*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 184-212, ISBN 978-83-67104-58-6.
72. Łasica W., Małek M., *Wpływ dodatku odpadowego pyłu polieteroeteroketonowego (PEEK) na wytrzymałość mechaniczną oraz parametry cieplne eko kompozytów cementowo-szklanych*, [w:] *Najnowsze doniesienia z inżynierii materiałowej*, Wydawnictwo Naukowe TYGIEL sp. z o.o., Lublin 2022, s. 213-238, ISBN 978-83-67104-58-6.
73. Abra O., Ftima M.B., *Strength reduction design method for reinforced concrete structures: Generalization*, *Engineering Structures*, 258, 2022, s. 114134.
74. Zhou M., Wu Z., Ouyang X., Hu X., Shi C., *Mixture design methods for ultra-high-performance concrete – a review*, *Cement and Concrete Composites*, 124, 2021, 104242.
75. Larrard F., *Ultrafine particles for the making of very high strength concretes*, *Cement and Concrete Research*, 19(2), 1989, s. 161-172.
76. Stovall T., Larrard F., Buil M., *Linear packing density model of grain mixtures*, *Powder Technology*, 48(1), 1986, s. 1-12.
77. Ji T., Chen C.-Y., Zhuang Y.-Z., Chen J.-F., *A mix proportion design method of manufactured sand concrete based on minimum paste theory*, *Construction and Building Materials*, 44, 2013, s. 422-426.
78. PN-EN 206+A2:2021-08, *Beton – Wymagania, właściwości, produkcja i zgodność*, Polski Komitet Normalizacyjny.
79. Gantner E., Chojczak W., *Materiały budowlane. Spoiwa, kruszywa, zaprawy, betony. Ćwiczenia laboratoryjne*, Oficyna wydawnicza Politechniki Warszawskiej, Warszawa 2013, ISBN 978-83-7814-020-7.
80. Jamróży Z., *Beton i jego technologie*, Wydawnictwo Naukowe PWN, 1, Warszawa 2021, ISBN 9788301182106, EAN:9788301182106.
81. Zieliński K., *Podstawy technologii betonu*, Wydawnictwo Politechniki Poznańskiej, Wydanie V, Poznań 2021, ISBN 978-83-7775-649-2.
82. Dondelewski H., Januszewski M., *Betony cementowe*, Wydawnictwo Naukowe PWN, 2019, EAN:9788301201678.
83. Januszewski M., *Beton towarowy*, Wydawnictwo Naukowe PWN, 2016, ISBN 9788301189013, EAN:9788301189013.
84. PN-EN 12390-2:2019-07, *Badania betonu – Część 2: Wykonywanie i pielęgnacja próbek do badań wytrzymałościowych*, Polski Komitet Normalizacyjny.
85. PN-EN 12390-1:2013-03, *Badania betonu. Część 1: Kształt, wymiary i inne wymagania dotyczące próbek do badań i form*, Polski Komitet Normalizacyjny.
86. PN-EN 12390-3:2019-07, *Badania betonu – Część 3: Wytrzymałość na ściskanie próbek do badań*, Polski Komitet Normalizacyjny.
87. PN-EN 12390-4:2020-03, *Badania betonu – Część 4: Wytrzymałość na ściskanie – Wymagania dla maszyn wytrzymałościowych*, Polski Komitet Normalizacyjny.
88. PN-EN 12390-13:2021-12, *Badania betonu – Część 13: Wyznaczenie siecznego modułu sprężystości przy ścisnaniu*, Polski Komitet Normalizacyjny.
89. Strona internetowa: <https://www.epsilontech.com/products/averaging-axial-extensometer-for-large-compression-samples-model-3542ra/> [data dostępu: 29.08.2022].
90. Strona internetowa: <https://www.epsilontech.com/products/diametral-rock-and-concrete-testing-extensometer-model-3975/> [data dostępu: 29.08.2022].
91. Strona internetowa: <https://www.epsilontech.com/> [data dostępu: 29.08.2022].

92. Strona internetowa: <https://www.epsilontech.com/products/rock-concrete-and-asphalt-extensometers/> [data dostępu: 29.08.2022].
93. Strona internetowa: <https://www.epsilontech.com/products/axial-torsional-extensometer-calibrator-model-3590at/> [data dostępu: 29.08.2022].
94. Zhu X., Gao Y., Dai Z., Corr D.J., Shah S.P., *Effect of interfacial transition zone on the Young's modulus of carbon nanofiber reinforced cement concrete*, Cement and Concrete Research, 107, 2018, s. 49-63.
95. Marques A.I., Morais J., Morais P., M. do Rosario Veiga, Santos C., Candeias P., Gomes Ferreira J., *Modulus of elasticity of mortars: Static and dynamic analyses*, Construction and Building Materials, 232, 2020, s. 117216.

Wpływ dodatku odpadowych wiórów stalowych na wartości modułu Young'a ekokompozytów cementowo-szklanych

Streszczenie

Temat zawarty w rozdziale dotyczy wpływu dodatku odpadowych wiórów stalowych na wartości stałej materiałowej w postaci modułu Young'a określanego dla stwardniałych ekokompozytów cementowo-szklanych. Badany wysokowytrzymały ekokompozyt poddano modyfikacji materiałowej dodatkiem odpadowych wiórów stalowych generowanych w procesie obróbki CNC półproduktów stalowych dla branży mechanicznej. Stalowy składnik odpadowy dozowano w odniesieniu do masy kruszywa, stanowił częściowy substytut kruszywa w składzie ekokompozytu. Badania w zakresie wytrzymałości charakterystycznej na ściskanie oraz wyznaczenie wartości modułu sprężystości podłużnej określono po 28 dniach dojrzewania ekokompozytów. Przedstawiono dwie metody badawcze, pozwalające określić wartości modułu Young'a, wykorzystujące zestaw ekstensometrów mierzących wartości przemieszczeń liniowych podłużnych i poprzecznych. Zarejestrowane wartości przemieszczeń liniowych umożliwiły wyznaczenie wartości odkształceń liniowych w kierunku podłużnym i poprzecznym próbek cylindrycznych. Scharakteryzowano składniki tworzące ekokompozyt cementowo-szklany, tj. spoiwo hydrauliczne specjalne, frakcje granulatów szkła sodowego, mączkę szklaną, płynne domieszki chemiczne oraz dodatek w postaci odpadowych wiórów stalowych. Przedstawiono metodę projektowania składu mieszanek cementowo-szklanych, promującą rozwiązania proekologiczne, tj. równoległe zastosowanie dwóch odpadowych spoiw przemysłowych oraz konstrukcje stosujące okruczowych złożonych z mączki i grup frakcji szkła sodowego.

Słowa kluczowe: ekokompozyt, moduł Young'a, cement, szkło, odpady stalowe

Influence of the addition of waste steel chips on the values of Young's modulus of eco-cement-glass composites

Abstract

The topic included in the chapter concerns the impact of the addition of waste steel chips on the value of the material constant in the form of The Youn's module determined for hardened cement-glass eco-composites. The tested high-strength eco-composite was subjected to material modification with the addition of waste steel chips generated in the CNC machining process of steel semi-finished products for the mechanical industry. The steel waste component was dosed in relation to the aggregate mass, it was a partial substitute for aggregate in the composition of the eco-composite. Tests in the field of characteristic compressive strength and determination of the value of the longitudinal elastic modulus were determined after 28 days of maturation of eco-composites. Two test methods were presented to determine the values of Young's modulus, using a set of extensometers measuring the values of longitudinal and transverse linear displacements. The recorded values of linear displacements made it possible to determine the values of linear deformations in the longitudinal and transverse direction of cylindrical samples. The components forming the eco-cement-glass composite, i.e. special hydraulic binder, sodium glass granulate fractions, glass meal, liquid chemical admixtures and an additive in the form of waste steel chips, were characterized. The method of designing the composition of cement-glass mixtures, promoting pro-ecological solutions, i.e. the parallel use of two waste industrial binders and the construction of crumb piles composed of meal and groups of soda glass fractions, was presented.

Keywords: eco-composite, Young's module, cement, glass, steel waste

Stanowisko spawalnicze ramy elektrycznego motocykla LEM Photon

1. Wstęp

Stanowisko spawalnicze służy do precyzyjnego spójczonowania elementów zaprojektowanej konstrukcji i ma umożliwić wykonanie dokładnych połączeń spawanych. Ewentualne błędy w wykonaniu stanowiska spawalniczego rzutują na samo wykonanie elementu. W artykule zostanie opisany proces realizacji stanowiska spawalniczego do ramy LEM Photon. LEM Photon to elektryczny motocykl stworzony na zawody Motostudent 2020 przez studentów Politechniki Wrocławskiej. Motostudent to międzynarodowe zawody przeznaczone dla studentów, które polegają na zaprojektowaniu i wykonaniu prototypu motocykla. Każdy motocykl przechodzi przez testy dynamiczne i statyczne, następnie zespoły, którym udało się zakwalifikować, biorą udział w wyścigu na międzynarodowym torze na Aragonie w Hiszpanii (5 okrążeń). Motocykl musi zostać zaprojektowany i wykonany w precyzyjny i przemyślany sposób. Oznacza to, że konstruktorzy muszą wziąć pod uwagę szereg czynników, takich jak:

- sztywność elementów (zarówno elementów spawanych, jak i elementów stanowiska spawalniczego);
- odebranie odpowiednio poszczególnych stopni swobody (zabezpieczonych przed przemieszczaniem);
- wpływ ciepła (wytwarzanego podczas procesu spawania) na odkształcenia ramy oraz stanowiska spawalniczego;
- montaż stanowiska (musi być możliwe złożenie stanowiska spawalniczego w warunkach warsztatowych);
- liczba zmian stanowiska spawalniczego (pożądane jest jak najmniejsza ilość przebrożeń stanowiska. Pozwala to zaoszczędzić czas i zmniejszyć ryzyko powstania błędów montażowych). Zmiany przy stanowisku spawalniczym oznaczają montaż dodatkowych elementów, takich jak ściany i elementy konstrukcji spawanej;
- odpowiednio dobrana technologia obróbki elementów stanowiska spawalniczego i elementów ramy pod montaż (obróbka plastyczna, cięcie laserowe, frezowanie i toczenie);
- odpowiednia przestrzeń do wykonywania połączeń spawanych dla spawacza;
- budżet na wykonanie takiej konstrukcji.

Ostatni podpunkt, czyli środki finansowe, jest ściśle powiązany z tym, że wykonanie ramy LEM Photon jest produkcją jednostkową. Oznacza to, że dane stanowisko spawalnicze, zostało przewidziane tylko pod wytworzenie jednej ramy. Po zakończeniu operacji stanowisko, zostało rozebrane, a niektóre części, takie jak: tulejki, śruby, podkładki, nakrętki – posłużyły do budowy innych projektów (reszta stanowiska, została zutylizowana). W związku z tym, że dane stanowisko jest produkcją jednostkową,

¹ aleksandermadralla@gmail.com, Wydział Mechaniczny, Politechnika Wroclawska.

² lentner.michal55@gmail.com, Wydział Mechaniczny, Politechnika Wroclawska.

przekłada się na to, że projektant nie musi przewidywać wielokrotnego montażu i demontażu stanowiska w jak najkrótszym czasie oraz czasu eksploatacji stanowiska, co z kolei powoduje obniżenie kosztów jego wyprodukowania [1].

Rama motocykla elektrycznego LEM Photon jest konstrukcją łączącą dużą wytrzymałość z małą masą. Jest to jeden z najbardziej odpowiedzialnych elementów konstrukcji. Każdy błąd w wykonaniu odbiegający od projektu będzie skutkował pogorszeniem parametrów motocykla (takich jak: sztywność, minimalna waga, tendencja do wpadania w wibracje, czy też trudność w montażu komponentów współpracujących [2]). Najbardziej kluczowymi miejscami do spozycjonowania była główka ramy, mocowanie osi wahacza, jego amortyzatora oraz kiwki. Precyzja wykonania połączeń w tych miejscach przekłada się na zachowanie odpowiedniej geometrii zawieszenia, a co za tym idzie właściwości jezdnych pojazdu.

2. Konstrukcja i założenia technologiczne

Po zaprojektowaniu ramy motocykla LEM Photon została przeprowadzona analiza sposobu jej wykonania. Uznano, że skonstruowanie stanowiska spawalniczego będzie najodpowiedniejszym rozwiązaniem. W oparciu o wyniki analizy, doświadczenie i wnioski wyciągnięte z realizacji projektu, sporządzona została lista wytycznych (funkcje i warunki, które muszą zostać spełnione), które zostały rozpatrzone pod względem technologicznym i finansowym. Celem stanowiska spawalniczego było uzyskanie jak najlepszej konstrukcji spawanej (jakości połączeń spawanych, dokładności geometrycznej oraz wymiarowej [3]). Poprzez stosowanie się do ustalonych wytycznych możliwe było uniknięcie popełnienia błędów konstrukcyjnych oraz uzyskanie pożądanego efektu końcowego. Poniżej opisano najważniejsze funkcje i wymagania stanowiska spawalniczego.

2.1. Pozycjonowanie elementów, modułowość i precyzja wykonania stanowiska

Pozycjonowanie elementów było ściśle związane z geometrią stanowiska. Do ustalenia elementów spawanych w odpowiedni sposób zostały użyte ściany, zamki technologiczne oraz tulejki. Wszystkie elementy wymagające współosiowości połączono jedną osią pomocniczą. Role takich osi pomocniczych pełniły: śruby pasowane i wytoczone specjalne elementy. W stanowisku istotne do spozycjonowania współosiowego były takie elementy, jak oś wahacza oraz główka ramy. Szczegółowo zostało to opisane i zobrazowane w rozdziale poświęconym właśnie tym elementom.

Realizacja spozycjonowania elementów (innych niż osiowe) polegała na przyjęciu strategii stworzenia stanowiska spawalniczego ramy składającego się z elementów modułowych, czyli składanych ścianek, podstaw oraz zamków technologicznych. Konstrukcję modułową wybrano m.in. ze względu na następujące zalety:

- wykorzystanie cięcia laserowego umożliwiło uzyskanie oczekiwanego kształtu, jednocześnie zachowując dużą precyzję wykonania wraz z uwzględnieniem minimalizacji kosztów. Przyjęty wariant to optymalna wersja dla tego typu produkcji jednostkowej. Podczas obróbki laserowej możliwe było dokładne uzyskanie (do wartości 0,1 mm) położenia środków otworów oraz rozmieszczenie pozostałych kształtów [4];

- możliwość projektowego uproszczenia konstrukcji. Dzięki zastosowaniu modułów było możliwe uproszczenie konstrukcji. Pozwoliło to na uniknięcie skomplikowanych procesów produkcyjnych (np. uwzględnienie gięcia elementów stanowiska), upraszczając tym samym zaprojektowanie stanowiska.

Elementy modułowe należało zaprojektować, zwracając jednocześnie uwagę na precyzję połączeń kształtowych. W związku z tym można było uzyskać połączenia zastosowane w stanowisku spawalniczym ramy LEM Photon, które wzajemnie niwelowały swoje błędy wymiarowe.

Należy także nadmienić, że elementy modułowe zostały zastosowane w celu uniknięcia procesu spawania. Jest to szczególnie ważne, ponieważ spawane elementy odkształcają się, co w rezultacie mogło negatywnie wpływać na pozycjonowanie oraz dokładność wykonania stanowiska.

2.2. Sztywność stanowiska spawalniczego

Sztywność samego stanowiska w dużym stopniu wpływała na finalną jakość wyrobu [5]. Jeśli stanowisko nie zostałoby odpowiednio pozbawione stopni swobody jego części nie byłyby stabilne (np. ścianki, które nie zostałyby prawidłowo przytwierdzone do podstawy miałyby możliwość przechylania się w otworze). W tym celu zastosowano odpowiednie podparcie ścianek i przytwierdzenie ich do podstawy z wykorzystaniem metod kasowania luzów (dokładniej zostanie to omówione w rozdziale poświęconym montażowi).

Ponadto sama sztywność wiąże się z zabezpieczeniem elementów przed przypadkowym przemieszczeniem (np. oparcie się o ściankę mogłoby zmienić pozycję samego elementu spawanego). W tym celu stosuje się większe siły montażowe podczas skręcania elementów stanowiska.

2.3. Zabezpieczenie elementów przed przemieszczeniem

Jak już zostało wspomniane, należało unieruchomić wszystkie elementy spawane przed przypadkowym przemieszczeniem (wyjątkiem były jedynie niektóre żebra). Było to wyjątkowo ważne, ponieważ podczas procesu, spawacz omyłkowo mógłby przesunąć elementy względem siebie. Mogłoby to skutkować nieprawidłową geometrią ramy (zmianą kątów, odległości lub połączeń). Tę funkcję przede wszystkim spełniały zamki technologiczne. W założeniach każdy okrągły profil miał być w miarę możliwości umocowany przynajmniej w dwóch miejscach. W przypadku giętej rury, zamki umieszczone były po dwóch stronach gięcia, co częściowo pozwoliło na zniwelowanie ewentualnych niedokładności procesu gięcia i ich wpływu na pozycję elementu. Zamki technologiczne między tulejami zamka a rurą były tak zaprojektowane, aby występowało ciasne pasowanie, co zapewniało zabezpieczenie przed przemieszczeniem i odkształceniami cieplnymi.

Skręcenie elementów modułowych z odpowiednią siłą korzystnie wpłynęło na zabezpieczenie ich przed przemieszczeniem.

2.4. Zabezpieczenie elementów przed odkształceniami cieplnymi

Zabezpieczenie przed przemieszczeniami na skutek odkształceń cieplnych było szczególnie ważną funkcją stanowiska z uwagi na możliwość zmiany kształtu i wystąpienia naprężeń resztkowych spowodowane działaniem temperatury [6]. Zastosowanie przynajmniej dwóch zamków technologicznych na jeden okrągły profil pozwoliło go

dobrze spozycjonować i zapobiec takim odkształceniom. Siły pochodzące od elementów pozycjonujących, wywołały wymagany moment tarcia, który zapobiegł przesuwananiu się elementów spawanych względem siebie oraz zredukował odkształcenia cieplne. Należy pamiętać, że użycie zbyt dużych sił montażowych może spowodować uszkodzenie elementu. W celu dodatkowego zabezpieczenia przed wpływem temperatury użyte zostały grube blachy i materiał odporny na działania wysokich temperatur, gdyż wysoka temperatura mogłaby z łatwością odkształcić cienkościenne elementy. Odkształcone w ten sposób stanowisko straciłoby swoją przydatność. Spoina ściąga względem siebie elementy, więc projekt musi uwzględniać, w którą stronę takie odkształcenie może zajść.

2.5. Wspomaganie procesu spawania

Stanowisko spawalnicze eliminuje w większości przypadków potrzebę ręcznego pozycjonowania elementów. Dzięki temu spawacz mógł skupić się wyłącznie na jakości powstających spoin. Projekt uwzględniał jak największy obszar dostępu wokół miejsc łączonych. Dostęp do problematycznych miejsc, został osiągnięty poprzez wykonanie dodatkowych otworów w podstawach stanowiska. Uwzględniając te otwory, rozplanowano ścianki i zamki technologiczne. Dodatkową cechą wspomagającą proces była możliwość obrócenia stanowiska (stanowisko spawalnicze ramy było symetryczne, więc posiadało dwie podstawy). Taka konstrukcja eliminowała potrzebę wykonywania spoin w pozycji pułapowej [7].

Najlepszym rozwiązaniem było złączenie większości elementów w stanowisku. Niestety, ze względu na skomplikowaną geometrię ramy nie wszystkie połączenia mogły zostać tak wykonane, dlatego część z nich jedynie szczepiono na stanowisku. Po wykonaniu spoin i szczepień wyjęto ramę ze stanowiska i przystąpiono do wykonania ciągłych spoin w miejscach szczepionych.

Warto nadmienić, że niektóre elementy, takie jak żebra w ramie, które nie były głównymi elementami nośnymi, zostały spozycjonowane i pospawane ręcznie. Nie wpłynęło to znacząco na ostateczną jakość wykonania ramy.

2.6. Inne aspekty, które powinien spełniać projekt stanowiska spawalniczego

Dołożone zostały starania, aby konstrukcja stanowiska była w miarę możliwości najprostsza i najtańsza w wykonaniu. Prostota wiązała się z wcześniej już poruszoną kwestią modułowości elementów, ilością użytych części i brakiem dodatkowych komplikacji w samym procesie montażu stanowiska spawalniczego ramy. Ze względu na konieczność wykonania dużej ilości otworów w stanowisku niosącego ewentualne ryzyko nieprawidłowego ich nagwintowania (np. pod kątem) podjęto decyzję o wykonaniu otworów przelotowych skręcanych od spodu podstawy nakrętkami.

Dodatkowo została dodana druga podstawa, aby uniknąć sytuacji, w której stanowisko opiera się na zakończeniach śrub, dzięki niej na śruby montażowe nie działa żadna dodatkowa siła osiowa, która mogłaby zmieniać położenie zamków technologicznych.

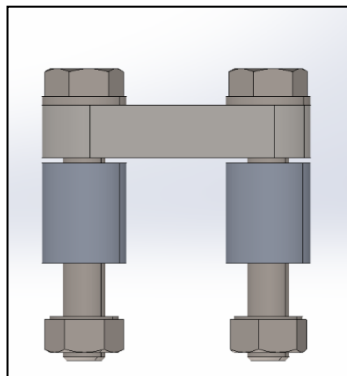
Materiał, który użyto do konstrukcji stanowiska, był łatwo obrabialny (obróbka laserowa), tani i wytrzymały na wysokie temperatury. Rolę tę dobrze spełniła stal konstrukcyjna S355, która przy jej znacznej grubości nie odkształciła się cieplnie.

3. Montaż stanowiska spawalniczego i jego elementów

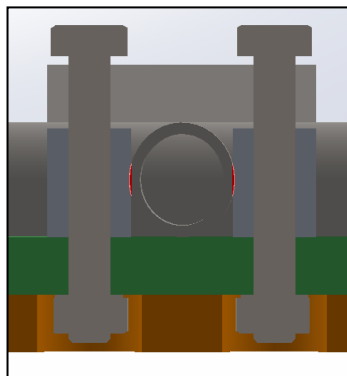
W celu wykonania dokładnych połączeń spawanych należy odebrać odpowiednio poszczególne stopnie swobody elementów łączonych. Elementy ramy LEM Photon zostały zabezpieczone przed przesunięciami i rotacją poprzez mocowanie profili okrągłych i blach stanowiska spawalniczego względem siebie.

3.1. Mocowanie profili okrągłych

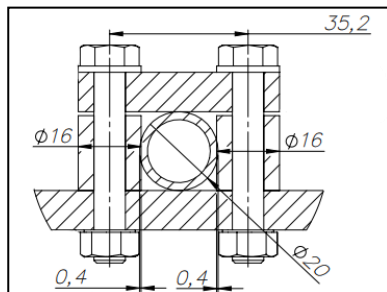
Profile okrągłe ramy zabezpieczono za pomocą zamków technologicznych, którego przykład, został pokazany na rysunku 1. Przykładowy zamek składa się z: tulejek, śrub, podkładek, nakrętek i górnej części, która łączy dwie śruby. Śruby są montowane w otworach wykonanych w podstawie stanowiska. Odległość pomiędzy środkami otworów, a co za tym idzie pomiędzy osiami śrub, zaprojektowana została, tak by profil okrągły był ciasno mocowany względem tulejek. Przed wykonaniem tego projektu eksperymentalnie wyznaczono, że optymalniejszą wartością wcisku jest 0,4 mm między jedną tulejką a profilem okrągłym (idea została zobrazowana na rysunku 2 i 3). Wcisk jest pożądany, ponieważ używając niewielkich sił wprasowania, rurki pomiędzy tulejki uzyskuje się jej stabilną pozycję.



Rysunek 1. Model obrazujący wcisk pomiędzy rurą i zamkiem technologicznym [opracowanie własne]

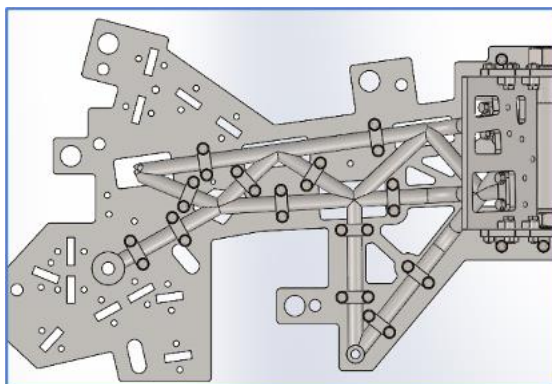


Rysunek 2. Model obrazujący wcisk pomiędzy rurą i zamkiem technologicznym [opracowanie własne]

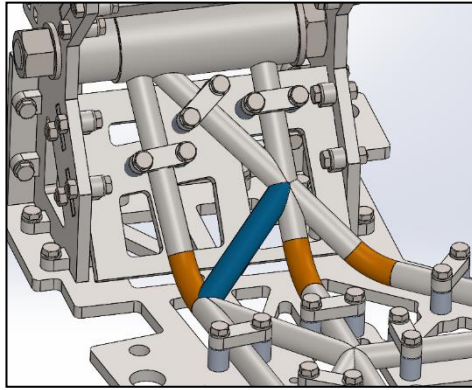


Rysunek 3. Fragment rysunku wykonawczego ramy LEM Photon [opracowanie własne]

Przy projektowaniu stanowiska należało również przewidzieć usztywnienie profilu okrągłego w przynajmniej dwóch miejscach (przykład rozlokowania zamków znajduje się na rysunku 4). Tam, gdzie jest to możliwe, została zastosowana większa ilość zamków technologicznych. Umieszczenie dwóch zamków trzymających jeden element nie zawsze jest wykonalne ze względu na: małe wymiary rurki, umiejscowienie rury w modelu oraz zostawienie miejsca na wykonanie spoiny. Przykładowe problematyczne miejsce, zostało pokazane na rysunku 5 (element wyróżniony kolorem niebieskim). Element ten blokowany był tylko przez inne już połączone i zablokowane rurki. Na rysunku 5 można również zwrócić uwagę na to, że profile okrągłe (elementy zaznaczone na pomarańczowo) nie zostały zablokowane w miejscach gięcia. Taka decyzja wynika z tego, że profil okrągły mógł zostać niedokładnie wygięty co utrudniłoby montaż. Taki zabieg pozwala na skompensowanie niedokładności gięcia [8].



Rysunek 4. Rzut z góry na złożenie stanowiska spawalniczego z ramą LEM Photon [opracowanie własne]



Rysunek 5. Fragment ramy LEM Photon w stanowisku spawalniczym. Na niebiesko zaznaczono rurę, która nie została zamocowana zamkami technologicznymi, na pomarańczowo zaznaczono rury, które nie zostały zamocowane w miejscach gięcia [opracowanie własne]

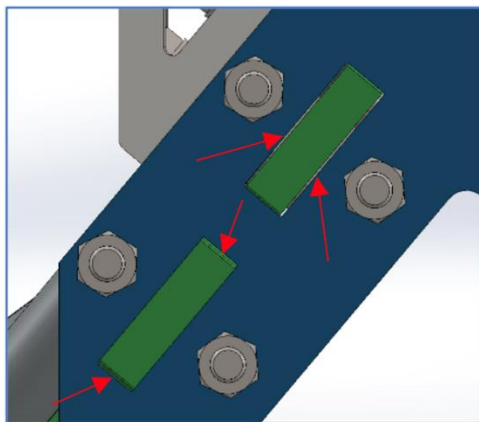
3.2. Mocowanie blach stanowiska spawalniczego względem siebie

Jednym z założeń przy projektowaniu stanowiska spawalniczego było takie rozmieszczenie elementów modułowych względem siebie w stanowisku, aby kompensowały wzajemnie błędy w swoim spozycjonowaniu. Równocześnie montaż takich elementów powinien być możliwy bez większych sił wprasowania. Stosując większe siły, może dojść do uszkodzenia elementów, co mogłoby przełożyć się na błędy w pozycjonowaniu ramy. Dlatego w konstrukcji były przewidziane luzy, które mają za zadanie ułatwienie montażu. Ponadto ścianka została wykonana w taki sposób, aby przy połączeniu jednej ścianki z drugą łączyły się w dwóch miejscach. Gdyby nie dokonano takiej operacji możliwy byłby ruch translacyjny prostopadły do wycięcia w ściance, w której jest mocowana druga ścianka. Kształt ścianki, o którym mowa, jest pokazany na rysunku 6.

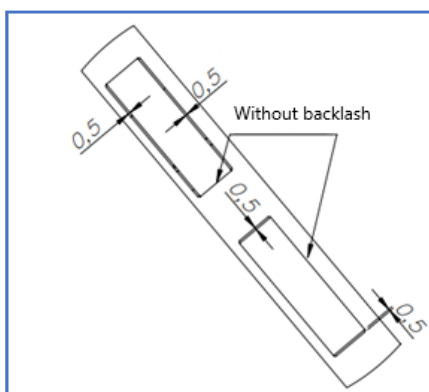


Rysunek 6. Ścianka za główką ramy, czerwone strzałki pokazują te elementy blachy, które łączą się z drugą ścianką [opracowanie własne]

W konstrukcji stanowiska występują dwa główne połączenia ścianek. Pierwszy, w którym ścianki nie były kontrolowane przez dodatkowe elementy. W takim przypadku luz występuje naprzemiennie (rys. 7 i 8).

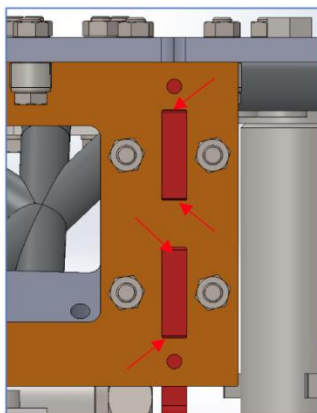


Rysunek 7. Przykład połączenia z luzem, czerwone strzałki sygnalizują luz [opracowanie własne]

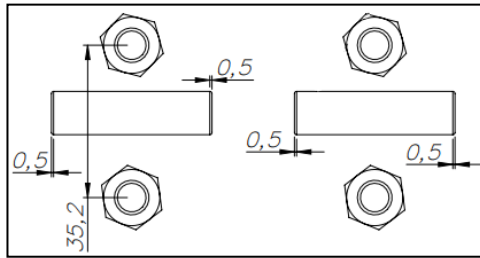


Rysunek 8. Fragment rysunku wykonawczego połączenia z luzem w dwóch kierunkach [opracowanie własne]

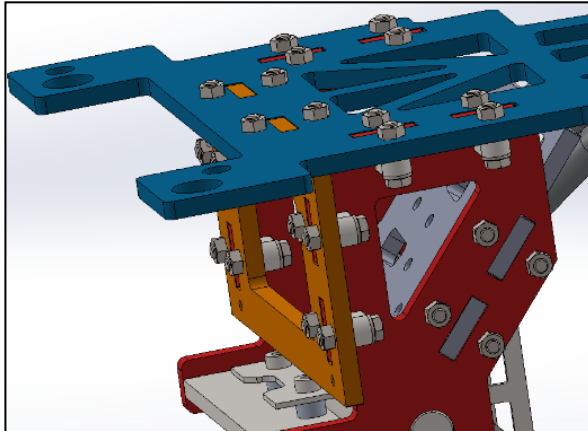
Drugi przypadek to taki, w którym występują dodatkowe elementy kontujące. Luz wtedy może być wykonany w tym samym kierunku (rys. 9 i 10). Na rysunku 11 pokazany jest omawiany przypadek. W tym wypadku dla ścianki pomarańczowej, ścianka czerwona i niebieska są elementami kontującymi.



Rysunek 9. Przykład połączenia z luzem w jednym kierunku, czerwona strzałka sygnalizuje luz [opracowanie własne]

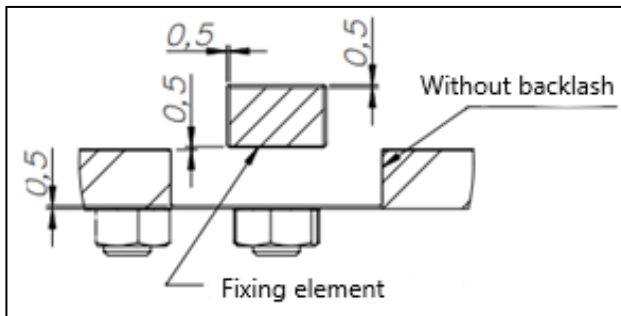


Rysunek 10. Fragment rysunku wykonawczego połączenia z luzem w jednym kierunku [opracowanie własne]

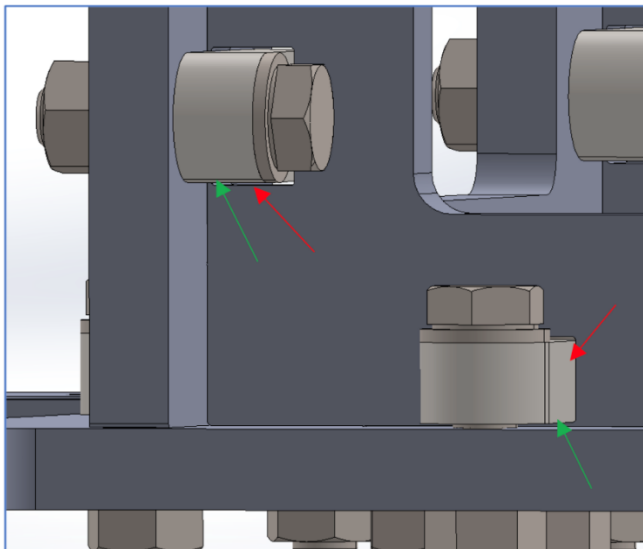


Rysunek 11. Fragment stanowiska spawalniczego, ścianki na czerwono i niebiesko kontrolują ściankę pomarańczową [opracowanie własne]

Otwory dla łącznika śrub od zamka technologicznego też powinny mieć przewidziany luz, w celu uniknięcia problemów montażowych (rys. 12 i 13). Zamek technologiczny ponadto powinien być tak przemyślany, aby łącznik dociskał ściankę do drugiej. W tym celu odsunięto łącznik od ścianki równoległej do niego tak, aby miał miejsce na lekkie odkształcenie (w zakresie sprężystym). Dzięki temu można zastosować większą siłę przy przykręcaniu śruby. To z kolei spowodowało dokładniejsze spozycjonowanie ścianek.

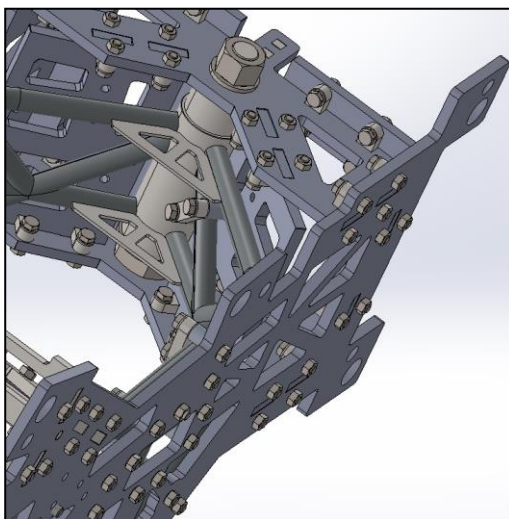


Rysunek 12. Fragment rysunku wykonawczego połączenia zamka technologicznego ze ścianką [opracowanie własne]

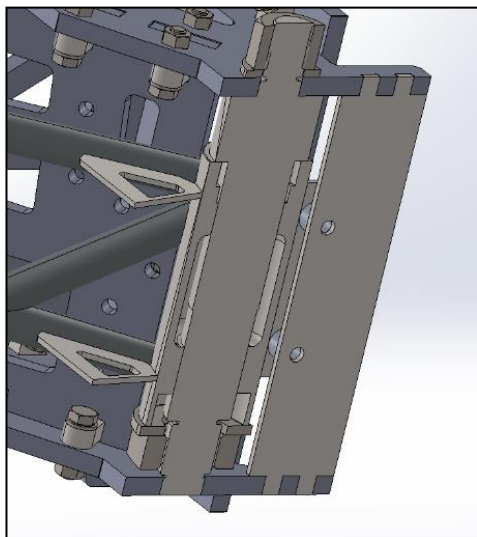


Rysunek 13. Fragment stanowiska spawalniczego. Przykład umiejscowienia zamka technologicznego. Czerwone strzałki wskazują poszerzone otwory o luz (0,5 mm), a zielone – przestrzeń od krawędzi otworu do drugiej ścianki [opracowanie własne]

Wszystkie elementy wymagające współosiowości, zostały połączone jedną osią pomocniczą. Ze względu na wymaganą dokładność, rolę osi pomocniczej pełni śruba pasowana lub elementy wytoczone specjalnie pod ten element ramy. Przykładem takiego elementu jest główka ramy (rys. 14). Końcówki osi pomocniczej, opierają się o otwory w ściankach stanowiska. Te elementy, które wymagają zabezpieczenia przed translacją, zabezpieczono odpowiednio nakrętką i specjalną tulejką wykonaną specjalnie do danego elementu (rys. 15).

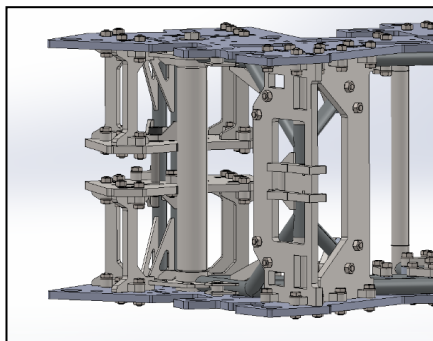


Rysunek 14. Fragment stanowiska spawalniczego – główka ramy [opracowanie własne]

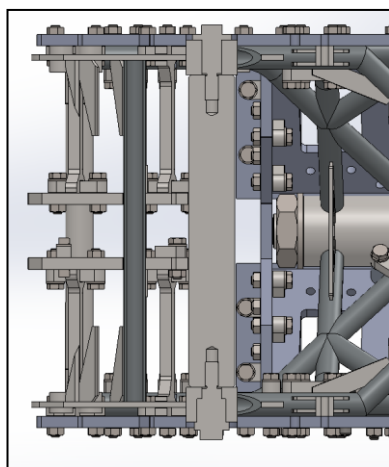


Rysunek 15. Fragment stanowiska spawalniczego – przekrój mocowania główki ramy [opracowanie własne]

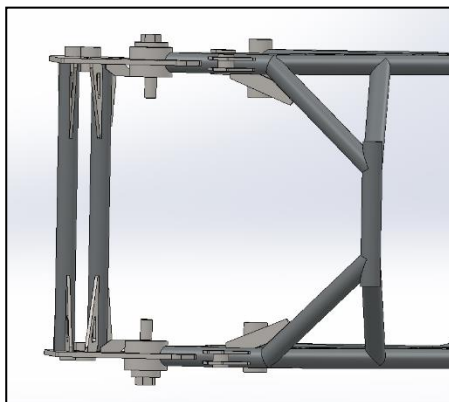
W podobny sposób została zabezpieczona oś wahacza (rys. 16-18).



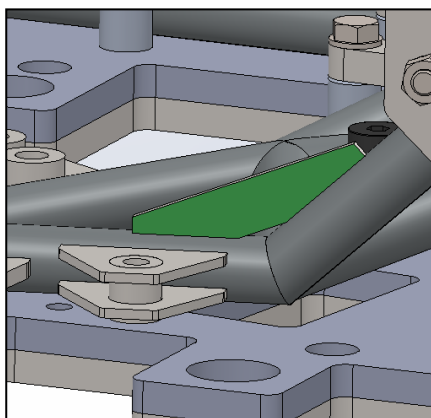
Rysunek 16. Fragment stanowiska spawalniczego – mocowanie osi wahacza [opracowanie własne]



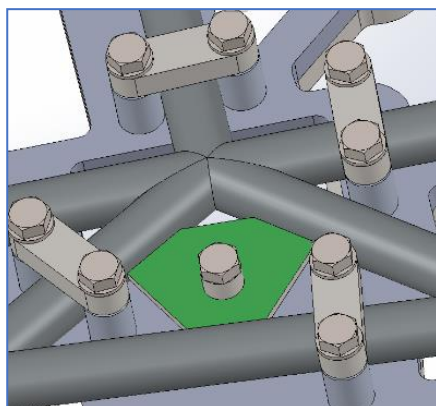
Rysunek 17. Przekrój sposobu mocowania osi wahacza [opracowanie własne]



Rysunek 18. Część osi wahacza z pokazanymi śrubami pasowanymi [opracowanie własne]



Rysunek 19. Żebro pozycjonowane ręcznie [opracowanie własne]



Rysunek 20. Żebro mocowane z dodatkowymi zabezpieczeniami przed przesunięciem lub rotacją [opracowanie własne]

Część żeber, te, u których nie jest wymagana wysoka dokładność położenia (rys. 19), spozycjonowano i zespawano ręcznie.

Żebra, od których wymagana była większa dokładność położenia, były zabezpieczone śrubą i blokowały się o już połączone rury (przykładowe zamontowanie żebra przedstawiono na rysunku 20).

4. Rezultaty, rozważania i wnioski

Zmontowane stanowisko wraz z elementami ramy gotowymi do spawania zostało przedstawione na rysunkach 21-23.



Rysunek 21. Połowa ramy umieszczona na stanowisku spawalniczym [opracowanie własne]

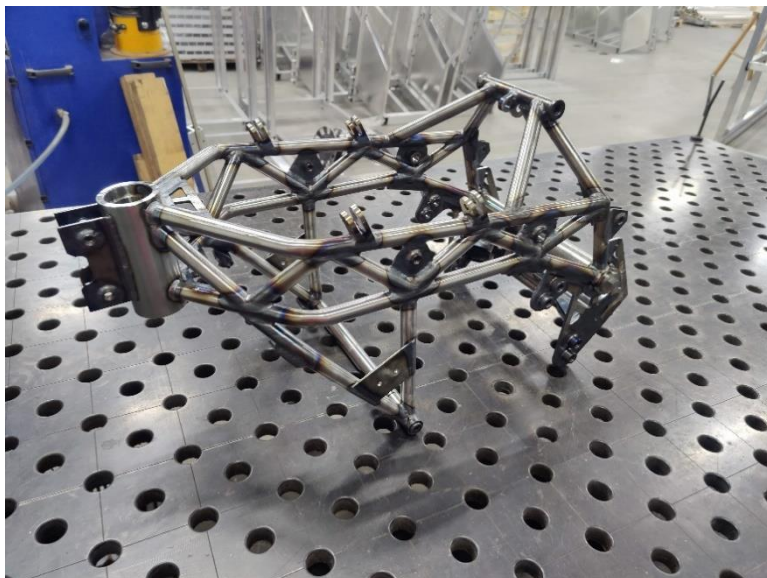


Rysunek 22. Złożone stanowisko z ramą gotową do spawania [opracowanie własne]



Rysunek 23. Złożone stawowisko z ramą gotową do spawania, widok z przodu [opracowanie własne]

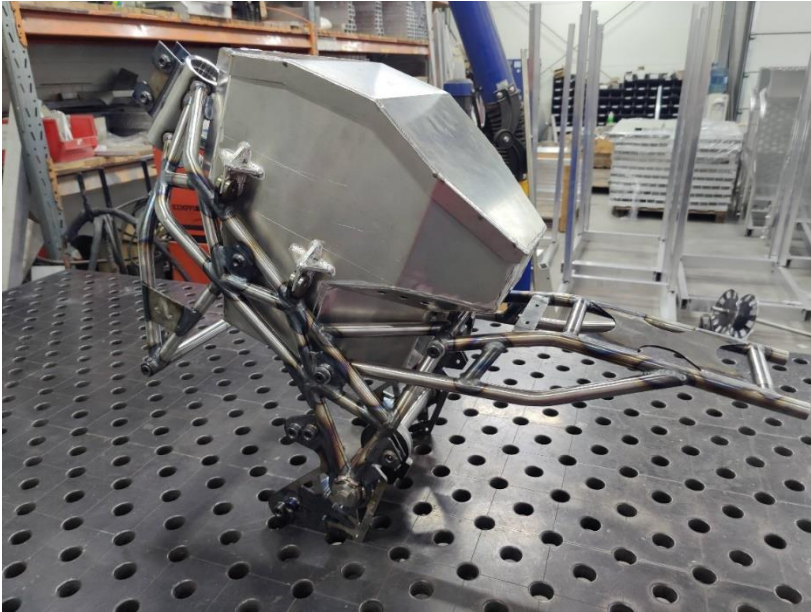
Gotowa konstrukcja nośna przedstawiona na rysunku 24 została później wykorzystana jako stawowisko do dokładnego spozycjonowania mocowań baterii wykonanej z aluminiowych blach. Cała rama wraz z baterią oraz subramą jako efekt końcowy wszystkich procesów została pokazana na rysunku 26. Dodatkowo na rysunku 25 przedstawiono przykładowy pospawany węzeł konstrukcyjny oraz połączenie z główką ramy.



Rysunek 24. Gotowa pospawana rama ze wszystkimi dodatkowymi mocowaniami [opracowanie własne]



Rysunek 25. Węzeł konstrukcyjny przy główce ramy [opracowanie własne]



Rysunek 26. Gotowa rama wraz z baterią oraz subramą [opracowanie własne]

Rama motocykla LEM Photon została zeskanowana przy użyciu ATOS Triple Scan S029 (z dokładnością do 0,1 mm). Skan został wykonany, aby wygenerować model 3D zespanej ramy do porównania z modelem CAD 14 [9]. Dzięki temu porównaniu było możliwe stworzenie modelu ramy z przedstawionymi odkształceniami.

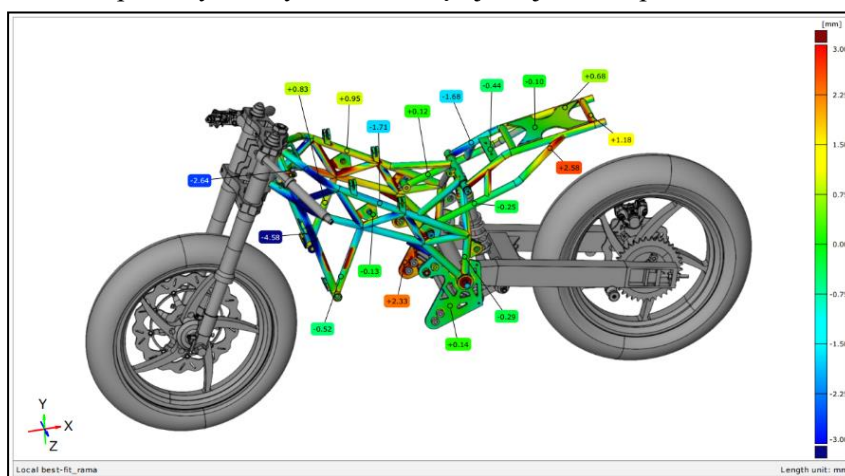
Rezultat przedstawiony na rysunku 27 był satysfakcjonujący, więc użycie stanowiska spawalniczego spełniło swoją funkcję. Jednakże wartości odkształceń muszą być przeanalizowane. Największe wartości wystąpiły w miejscu mocowania wahacza. Prawdopodobne jest, że grubość tych mocowań sprawiła występowanie tam największych temperatur przez wydłużony czas spawania tego obszaru.

Następnym miejscem generującym największe odkształcenia było mocowanie silnika. Wystąpiło to jako konsekwencja użycia na ten element blachy, która mogła łatwo zmieniać kształt pod wpływem temperatury łuku spawalniczego. Te odkształcenia zostały przewidziane przez konstruktora oraz ich wartości są akceptowalne i nie wpłynęły na osiągi motocykla. Można również łatwo zauważyć, że najdłuższe z rur mają największe wartości odkształceń.

Największe wartości odkształceń również występują wokół połączeń spawanych. Złącza były najbardziej narażone na działanie temperatury, więc w tych miejscach doszło do spiętrzenia naprężeń. Inną przyczyną takiego zjawiska mogły być trudności przy procesie spawania, dokładniej dojścia do wszystkich miejsc wokół złączy końcówką elektrody wolframowej TIGa, co jest niezbędne dla uzyskania dużej dokładności.

Często przed rozpoczęciem procesu spawania, elementy łączone zostają wstępnie nagrzane, aby zminimalizować gradient temperatur i osiągnąć najlepszy rezultat. Niestety w tym przypadku nie było możliwości takiego przygotowania materiału, więc spawacz musiał zwiększyć natężenie prądu, co wpłynęło na dokładność wykonania [10].

Pomimo nieznacznych odstępstw geometrycznych od modeli CAD rama została poprawnie pospawana, więc można stwierdzić, że stanowisko spawalnicze spełniło swoją rolę. Przed rozplanowaniem wykonania kolejnego stanowiska należy rozwiązać wyżej wspomniane problemy. W tym przypadku można było ich uniknąć poprzez zwiększenie temperatury lub wydzielenie więcej miejsca dla spawacza.



Rysunek 27. Wynik wykonania ramy motocykla na stanowisku spawalniczym [opracowanie własne]

Literatura

1. Oprzyrządowanie stanowiska spawalniczego: <https://patmar-eng.pl/przyrzady-spawalnice-montazowe> [data dostępu: 04.01.2023].
2. Hieu L.-T., Khoa N.X., Lim O., *An Investigation on the Effects of Input Parameters on the Dynamic and Electric Consumption of Electric Motorcycles*, Sustainability, 13, 2021, s. 72-85.
3. Gandhe E., *Influence of GMAW Process Parameters and Selection Techniques on the Quality of a Welded Joint*, 7(12), 2019, s. 741-758.
4. Sawada K., Hara T., *Microstructure characterization of heat affected zone after welding in Mod.9Cr-1Mo steel*, 101, 2015, s. 106-113.

5. Ehlers T., Lachmayer R., *Design of a Motorcycle Triple Clamp Optimised for Stiffness and Damping*, [w:] Pfingstl S., Horoschenkoff A., Höfer P., Zimmermann M. (red.), *Proceedings of the Munich Symposium on Lightweight Design 2020*, Springer Vieweg, Berlin 2021, s. 1-17.
6. Steen W.M., Mazumder J., *Laser Material Processing*, Springer London, Londyn 2010, s. 1-558.
7. Cho D.W., Na S.J., Cho M.H., Lee J.S., *A study on V-groove GMAW for various welding positions*, *Journal of Materials Processing Technology*, 213(9), 2013, s. 1640-1652.
8. Kozlov A., Khaliulin E., Platov S., Nalimova M., Nekt V., *Evaluation of deformations of thin-walled pipes during cold bending with rolling*, *Letters on Materials*, 9(2), 2019, s. 218-222.
9. Shirakawa K., Kobashi S., Kurono Y., Shono M., Isaji O., *3D-Scan Millimeter-Wave Radar for Automotive Application*, *FUJITSU TEN TECH*, 38, 2013, s. 1-5.
10. Saraev Y.N., Golikov N.I., Maksimova E.M., Sidorov M.M., *Influence of low temperature on structure and impact strength of structural steels welded joints*, *Journal of Physics: Conference Series*, 1546(1), 2020.

Stanowisko spawalnicze ramy elektrycznego motocykla LEM Photon

Streszczenie

Dokładność wykonania ramy motocykla jest kluczowym elementem. Zapewnia ona bezpieczeństwo kierowcy i zwiększenie osiągnięć motocykla. W celu odpowiedniego wykonania połączeń spawanych w ramie motocykla, elementy muszą być odpowiednio zabezpieczone. Stanowisko spawalnicze ma na celu zabezpieczyć pozycję elementów i usztywnić konstrukcję. Poniższy artykuł przedstawia założenia konstrukcyjne, technologiczne oraz realizację stanowiska spawalniczego ramy motocykla LEM Photon.

Słowa kluczowe: stanowisko spawalnicze, motocykl elektryczny, spawanie ramy

Welding stance of frame electric motorcycle LEM Photon

Abstract

The manufacturing accuracy of the motorcycle's frame is very important. It connects directly with the safety of the rider and the performance of the motorcycle. In order to weld the frame, the elements must be properly secured. The construction which provides the stiffness and the security is welding stance. The following article show realization of welding stance for frame LEM Photon and the construction and technological assumptions

Keywords: welding stance, electric motorcycle, frame welding

Niedeterministyczny algorytm zachłanny do zastosowań w strategii marszrutyzacji – studium przypadku

1. Wprowadzenie

Opracowanie zautomatyzowanego komponentu informatycznego ułatwiającego rozplanowanie dostaw towarów do zadanych punktów odbioru (destynacji) jest jedną z podstawowych potrzeb sygnalizowanych przez rynek TLS (ang. *Transport, Logistic, Spedition*). Umiejętne przyporządkowanie zadań do zasobów często uzależnione jest od kompetencji i doświadczenia planistów i logistyków. Istotną kwestią, obok minimalizacji uogólnionego kosztu rozwiązania, ale także zachowanie jego dopuszczalności, tj. akceptowalność z uwzględnieniem różnego typu warunków i ograniczeń. Innym wymogiem jest jak najniższa czasochłonność procesu planowania, istotna szczególnie w przypadku konieczności dokonywania modyfikacji rozwiązania związanej z replanowaniem i doplanowaniem dostaw.

Wdrożenie skutecznego modelu zarządzania zasobami transportowymi, pozwalającego na jego efektywne wykorzystanie wiąże się z koniecznością opracowania i wdrożenia odpowiednich modeli obliczeniowych i algorytmów wspomagających proces planowania i optymalizacji obsługi zamówień. Zagadnienia te łączą się z problematyką zarządzania zasobami, takimi jak *Work Force Management* (WFM) [1] czy *Enterprise Resource Planning* (ERP) [2] z szeroką problematyką zagadnień transportowych. Są to zazwyczaj tzw. problemy NP-trudne, do których między innymi zalicza się problem marszrutyzacji (ang. *Vehicle Routing Problem – VRP*) [3, 4], problemy planowania kadry *Nurse Scheduling* (NSP) [5] oraz *Job Shop Scheduling* (JSP) [6, 7], problem plecakowy (ang. *knapsack problem*) [8], czy problem przypisywania w grafie dwudzielnym (ang. *assignment problem*) [9]. W przypadku zarządzania flotami transportowymi szczególnie istotny jest problem marszrutyzacji. Wspólnym wyzwaniem dla problemów NP-trudnych jest dobór adekwatnej metody optymalizacji, charakteryzującej się skalowalnością i uniwersalnością zastosowania. Dla problemów małej klasy rekomendować można dokładne metody programowania dynamicznego, takie jak algorytm Bellmana-Heldak-Karpa [10]. Jego złożoność obliczeniowa szacowana na $O(n^2 2^n)$ oraz wymagania przestrzenne (pamięciowe) rzędu $n \cdot 2^n$ w praktyce ograniczają jego realne zastosowanie dla instancji problemu VRP o liczbie destynacji n nie większym niż 30-40. Tymczasem, w wielu przypadkach wymagana jest obsługa kilkuset i więcej destynacji. W takich przypadkach rekomenduje się stosowanie heurystycznych metod przybliżonych.

Algorytmem, na bazie którego można opracowywać szybkie i akceptowalnie wydajne metody planowania, jest algorytm zachłanny (ang. *Greedy algorithm*) [11]. W najprostszych przypadkach, takich jak problem komiwojażera (ang. *Travel Salesman Problem – TSP*), sprowadza się on do poszukiwania rozwiązania na zasadzie strategii „najbliższego sąsiada”. Jednakże w bardziej zaawansowanych wariantach problemów

¹ r.belka@tu.kielce.pl, Katedra Systemów Informatycznych, Wydział Elektrotechniki, Automatyki i Informatyki, Politechnika Świętokrzyska, www.tu.kielce.pl.

marszrutyzacji, implementacja algorytmu musi ulec pewnym modyfikacjom – dotyczy to w szczególności planowania zadań dla flot heterogenicznych z dodatkowymi ograniczeniami. Algorytm musi bowiem uwzględniać niedopuszczalność określonych rozwiązań, spowodowaną nałożonymi warunkami. Analizy zachowań tego typu algorytmu dla złożonego i rzeczywistego problemu marszrutyzacji pozwalają na obserwację silnego związku pomiędzy stopniem nałożonych ograniczeń o zwracanym rezultatem.

1.1. Algorytmy zachłanne w zastosowaniach optymalizacji strategii marszrutyzacji – przegląd literatury

Algorytm zachłanny postępuje zgodnie z heurystyką rozwiązywania problemów, polegającą na dokonywaniu wyboru, który jest lokalnie optymalny na aktualnym etapie [11]. Rozwiązanie zachłanne jest zawsze optymalne dla pewnych popularnych typów problemów kombinatorycznych, takich jak znajdowanie minimalnego drzewa rozpinającego (*Minimum Spanning Tree*), czy znajdowanie najkrótszej ścieżki (ang. *Shortest Path Problem*). W przypadku innych zagadnień kombinatorycznych, wliczając w to VRP, strategia zachłanna nie daje rozwiązania optymalnego, jednak metoda ta może dostarczyć w rozsądnym czasie przybliżone rozwiązania, które mogą stanowić punkt wyjścia dla opcjonalnej reoptymalizacji. Z tego względu algorytm zachłanny wciąż stanowić może ważny komponent bardziej zaawansowanych algorytmów o strukturze hybrydowej lub wieloetapowej. W literaturze z ostatnich lat znaleźć można liczne przykłady zastosowań algorytmu zachłannego w rozwiązywaniu określonych wariantów problemu marszrutyzacji. Przykładowe dyskusje na temat zastosowania heurystyki „najbliższego sąsiada” (ang. *nearest neighbour*), stanowiącej uproszczony przypadek zachłanności do rozwiązania klasycznego problemu VRP z ograniczeniem na ładowność (ang. *Capacitated VRP – CVRP*), znaleźć można w pracy [12]. Autorzy porównywali tam dwa zbliżone algorytmy, z których pierwszy traktował rozwiązanie końcowe jako fragmenty grafu hamiltonowskiego znalezione metodą zachłanną dla problemu komiwojażera. Okazało się, że statystycznie korzystniejsze rozwiązanie zwraca algorytm B, w którym heurystyka „najbliższego sąsiada” aplikowana jest dla każdej z tras oddzielnie. Autorzy pracy [13] użyli algorytmu zachłannego do rozwiązania problemu CVRP, z równoczesnym uwzględnieniem dwuwymiarowego problemu pakowania (ang. *bin packing problem*). W pracy [14] rozwiązywano problem marszrutyzacji na przykładzie heterogenicznej floty statków. Nowością jest zorientowanie funkcji celu raczej na koszty eksploatacyjne niż na tradycyjnie rozważany dystans. Wynika to z faktu, że pojazdy transportujące są także zróżnicowane pod kątem kosztu jednostkowego. Autorzy nie uwzględniali jednak limitu DWL. W pracy [15] autorzy porównywali skuteczność algorytmu zachłannego z bardziej zaawansowaną metaheurystyką opartą o przeszukiwanie tabu (ang. *Tabu Search – TS*). Wyniki prezentowane w tej pracy sugerują, że rozwiązania zwracane przez prostą heurystykę „najbliższego sąsiada” charakteryzują się kosztem, o co najmniej 10% wyższym niż uzyskane za pomocą metody TS. Potwierdza to ogólne spostrzeżenia o „nieoptymalności” heurystyki zachłannej. Z kolei w pracy [16] rozwiązywano problem VRP w wariacie z jednoczesnym dowozem/odbiorem towarów (ang. *Pickup-Delivery*, VRP). W tym przypadku jednak algorytm zachłanny stanowił jedynie element służący generacji populacji startowej dla bardziej zaawansowanej metody metaheurystycznej realizującej optymalizację dwukryterialną obejmującą zarówno koszt dystansu, jak i koszt (karę) związaną z odchyleniem od założonych okien czasowych dostaw.

1.2. Motywacje, cel i zakres pracy

Rozwiązania problemów planistycznych znajdowane są w oparciu o podejście heurystyczne, bazujące na regionalizacji punktów dostaw, czy podejściu „zachłannym”. Uzyskane wyniki są wystarczająco korzystne w przypadku statycznej listy punktów dostaw, mogą jednak daleko odbiegać od rozwiązań optymalnych, w szczególności częstych zmian w rozkładzie punktów dostaw czy niewielkiego doświadczenia planisty. Tymczasem łatwo udowodnić, że w przypadku przeciętnej floty transportowej liczącej ok. 50 pojazdów, realizującej trasy o łącznym kilometrażu przekraczających 200 tys. km miesięcznie i zużywającym ok. 50 tys. litrów paliwa, nawet kilkuprocentowy uzysk wiąże się z istotnymi korzyściami finansowymi, przy okazji pozytywnie wpływając na poziom emisji gazów cieplarnianych. W celu efektywnego zarządzania zasobami transportowymi opracowano nowatorski system informatyczny o nazwie TARGET, który w skalowalny sposób integruje szereg metod optymalizacji problemów transportowych, w tym również prostą heurystykę opartą o algorytm zachłanny. W podstawowym założeniu, algorytm *Greedy* stanowi pierwszą fazę procesu obliczeniowego, którego wynik stanowi punkt wyjścia dla faz kolejnych. W efekcie prowadzonych prac zwrócono uwagę na silną zależność pomiędzy nasileniem ograniczeń na uzyskiwany wynik (koszt) końcowy, obserwując przy tym liczne anomalie w tym zakresie. Dlatego też w niniejszej pracy skupiono się na obserwacji zachowań algorytmu w obecności stochastycznych zaburzeń.

Stwierdzono, że w literaturze brak jest implementacji i analiz wariantów algorytmu zachłannego dla problemów klasy VRP, który cechowałby się niedeterministycznością, tj. zwracał odmienne (alternatywne) wyniki dla tego samego stanu zbioru zamówień i pojazdów. W związku z powyższym na potrzeby badań przeprowadzono modelową implementację algorytmu *Greedy* dla szybkiego planowania rozwoju towarów dla ~200 punktów dostaw. Reprezentantem zasobów była heterogeniczna flota transportowa złożona z 3 typów pojazdów różniących się ładownością i gabarytami. Dodatkowo, na wybrane destynacje nałożono ograniczenia na dojazd dla preferowanych zazwyczaj pojazdów o większych gabarytach. Wariant ten nazwano HFVRP-DWL (ang. *Heterogeneous Fleet Vehicle Routing Problem with Driveway Limitation*). Zaimplementowano i przeanalizowano klasycznie rozumiany algorytm zachłanny, zwracający jedno deterministyczne rozwiązanie dla zadanej listy zamówień. Następnie przeprowadzono badania i analizy zachowań algorytmu, wprowadzając stochastyczne zaburzenia do macierzy interdystansów. Badania prowadzono pod kątem metody generacji dopuszczalnej populacji startowej dla bardziej zaawansowanych i czasochłonnych probabilistycznych algorytmów ewolucyjnych.

2. Opis i założenia problemu optymalizacyjnego

2.1. Ogólna charakterystyka wariantu marszrutyzacji z flotą heterogeniczną i limitem na podjazd

Wariant marszrutyzacji z flotą heterogeniczną i limitem na podjazd (*HFVRP-DWL*) stanowi modyfikację klasycznego problemu VRP, w którym zakłada się istnienie zbioru destynacji reprezentowane przez graf skierowany, $G = (V, A)$, gdzie $V = \{1..D, D + 1..N + D\}$ jest zbiorem $D + N$ węzłów, przy czym D oznacza liczbę tzw. depotów, N – liczbę destynacji, zaś A oznacza zbiór ważonych krawędzi grafu. W rozważanym

przypadku przyjęto $D = 1$, a także założono kompletność grafu. Wagi jego krawędzi $d(i,j)$ reprezentują wartości dystansów pomiędzy poszczególnymi parami destynacji n_i i n_j . Wagi krawędzi mogą zostać przedstawione za pomocą macierzy kwadratowej o rozmiarze $N + D$, zwanej dalej tablicą interdystansów i oznaczanej jako $DistMx$. Element $DistMx[i,j]$ reprezentuje odległość pomiędzy destynacjami i oraz j , a jego wartość jest zazwyczaj głównym elementem, w oparciu o który wyznacza się funkcję kosztu wynikowego. Każdy z węzłów reprezentuje destynację, która zgłasza zapotrzebowania na usługi polegające na odbiorze pewnej ilości towaru. Żądanie dostarczenia towaru przez danego klienta (daną destynację) reprezentowane jest zazwyczaj przez ilość homogenicznego towaru wyrażonego w jednostkach umownych. W klasycznym wariacie problemu marszrutyzacji zakłada się, że pojedyncza usługa dostawy/odbioru towaru może być obsłużona w całości przez jeden pojazd, co nie zawsze jest prawdziwe w rzeczywistości. Należy zatem założyć, że zamówienia mogą być większe, zatem istnieje konieczność ich podziału. W rezultacie dana destynacja (klient) może być obsłużony przez więcej niż jeden pojazd, czyli podstawowe założenia klasycznego problemu VRP mogą być łamane.

W wariacie problemu marszrutyzacji z flotą heterogeniczną, zakłada się istnienie T typów pojazdów, różniących się ładownością i kosztem utrzymania. Niech $\forall t \in \{1..T\} \mathbb{V}max_t$ oraz $\mathbb{M}max_t$ oznaczają odpowiednio maksymalną ładowność objętościową i wagową pojazdu danego typu. Niech \mathbb{K}_t oznacza ilość dostępnych pojazdów typu t . Wówczas $K_{tot} = \sum_{t \in \{1..T\}} \mathbb{K}_t$ oznacza całkowitą liczbę pojazdów będących w dyspozycji, z kolei $\mathbb{V}_{tot} = \sum_{t \in \{1..T\}} (\mathbb{K}_t \cdot \mathbb{V}max_t)$ oraz $\mathbb{M}_{tot} = \sum_{t \in \{1..T\}} (\mathbb{K}_t \cdot \mathbb{M}max_t)$ reprezentować będą całkowitą objętość i masę przewożoną floty oraz stanowią jedną z podstawowych parametrów przestrzeni zasobów. Ponadto $\mathbb{C}_{M_t}, \mathbb{C}_{URF_t}$ oraz \mathbb{C}_{URV_t} oznaczają odpowiednio stałe koszty utrzymania ($M - Maintenance$), stałe koszty jazdy (URF – ang. *Unit Route Fix*) oraz zmienne koszty jazdy (URV – ang. *Unit Route Var*). Koszty \mathbb{C}_{M_t} reprezentują koszty utrzymania pojazdu, niezależnie czy jest on w trasie czy nie. Pojazdy większe mogą charakteryzować się wyższymi kosztami związanymi np. z garażowaniem, opłatami ubezpieczeniowymi itp. Koszty \mathbb{C}_{RF_t} utożsamiane są z jednostkowym (kilometrowym) kosztem przejazdu pojazdu pustego (bez ładunku). Z kolei \mathbb{C}_{RV_t} to koszty reprezentujące dodatkowe spalanie związane z przewożonym ładunkiem. Wówczas sumaryczny koszt jednostkowy dla odcinka trasy pomiędzy destynacją i a destynację j wyrażać się będzie prostą zależnością:

$$\mathbb{C}_{UR}[i,j] = \mathbb{C}_{URF} + \mathbb{C}_{URV} \cdot LoadMass[i,j] \quad (1)$$

gdzie $LoadMass [i,j]$ jest aktualną wagą przewożonego ładunku na tej trasie. Następnie koszt przemieszczenia się po krawędzi (i,j) grafu G można z kolei przedstawić następująco:

$$\mathbb{C}_R[i,j] = \mathbb{C}_{UR}[i,j] \cdot d[i,j] \quad (2)$$

Niech $\mathbb{R}^{(k \in \mathbb{K}_t)}$ oznacza trasę przebytą przez jeden z k pojazdów należących do typu t . Trasa ta reprezentowana jest przez sekwencję $h = \{1..H\}$ wierzchołków i (destynacji), przy czym i_1 oraz i_H oznaczają jeden z depotów. Wówczas wyrażenie (3):

$$\mathbb{C}^{(k \in \mathbb{K}_t)} = \mathbb{C}_{M_t} + \sum_{h \in \mathbb{R}^{(k \in \mathbb{K}_t)}} \mathbb{C}_R[i_h, j_{h+1}] \quad (3)$$

oznacza koszt całkowity tej trasy. Jeśli trasa $\mathbb{R}^{(k \in \mathbb{K}_t)}$ zawiera tylko 2 elementy $i_1 = i_2 \in D$, wówczas trasa taka jest trasą pustą, a jej koszt równa się \mathbb{C}_{M_t} . Oznacza to, że zaproponowane rozwiązanie nie przewiduje wykorzystania pojazdu k w redystrybucji towaru. Koszt całkowity jest zwykłą sumą kosztów $\mathbb{C}^{(k \in \mathbb{K}_t)}$.

$$\mathbb{C}_{tot} = \sum_{t \in T, k \in \mathbb{K}_t} \mathbb{C}^{(k \in \mathbb{K}_t)} \quad (4)$$

Jednocześnie wyrażenie (4) może być potraktowane jako funkcja celu. Warto zauważyć, że gdy $\mathbb{C}_{UR} = 1$, zaś $\mathbb{C}_{M_t} = 0$, powyższa funkcja kosztu reprezentuje długość trasy w kilometrach. W rozważanym przypadku uznano za zasadne użycie uproszczonej wersji funkcji kosztu z uwagi na cel, jakim jest uzyskiwanie jedynie przybliżonych rozwiązań, a także z powodu łatwiejszej możliwości porównania wyników z danymi rzeczywistymi.

Poszukiwanie jak najmniejszej wartości kosztu \mathbb{C}_{tot} jest podstawowym kryterium optymalności i celem algorytmu optymalizacyjnego. Przy czym spełnione muszą być przede wszystkim podstawowe warunki związane z ograniczoną ładownością poszczególnych towarów:

$$\forall_{t \in \{1..T\}, (k \in \mathbb{K}_t)} \sum_{h=2}^{|\mathbb{R}^{(k \in \mathbb{K}_t)}|-1} Vol_{i_h} \leq \mathbb{V}max_{t_k} \quad (5a)$$

$$\forall_{t \in \{1..T\}, (k \in \mathbb{K}_t)} \sum_{h=2}^{|\mathbb{R}^{(k \in \mathbb{K}_t)}|-1} Mass_{i_h} \leq \mathbb{M}max_{t_k} \quad (5b)$$

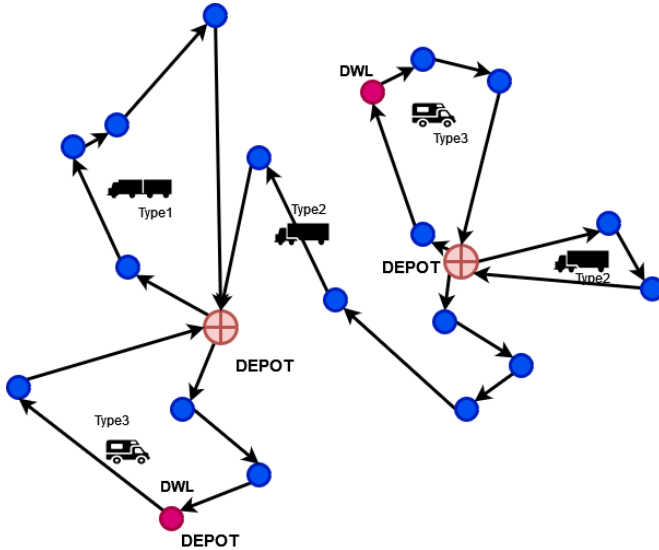
Przy czym pod wielkościami Vol i $Mass$ kryją się ilości towaru dostarczanych (odbieranych) z poszczególnych punktów znajdujących się na trasie przejazdu. Trasy, które spełniają powyższe warunki są trasami akceptowalnymi, tym samym wchodzą w skład rozwiązań dopuszczalnych.

Zaprezentowany problem, zakładający istnienie jednego depotu, można zaliczyć do wspomnianego wariantu *HFVRP-DWL*. W odróżnieniu od innych wariantów, wynik działania może zależeć od bazy dostępnych pojazdów, reprezentowanych przez wektor \mathbb{K} . W przypadku, gdy \mathbb{K}_t są nieograniczone, problem sprowadza się do bardziej klasycznej wersji, przy czym wybór typu pojazdu jest uzależniony od wzajemnej relacji \mathbb{C}_{M_t} , \mathbb{C}_{URF_t} oraz \mathbb{C}_{URV_t} . W praktyce zbiór jest ograniczony, wówczas należy spodziewać się różnych rezultatów w zależności od postaci \mathbb{K} . Utrudnia to obiektywną ocenę skuteczności algorytmu np. poprzez porównanie go z danymi rzeczywistymi.

Na rysunku 1 przedstawiono symboliczną wizualizację tras pojazdów, uwzględniającą wariant z dwoma depotami, flotą heterogeniczną oraz ograniczeniami na FDM. Występują tu 3 typy pojazdów. Dwie z destynacji odznaczają się ograniczeniami na dojazd wyłącznie pojazdami typu 3. Rozwiązanie zakłada, że trasa, która zawiera destynacje z ograniczeniami typu DWL muszą być obsługiwane przez pojazdy właściwego typu. Zauważmy przy tym, że problem DWL może być rozważany jedynie w wariacie VRP z flotą heterogeniczną.

W omawianym modelu założono, że każda z N destynacji musi posiadać informacje o dopuszczalności pojazdu typu t , ze względu na podjazd. Wówczas ze zbioru liczb $\{1..T\}$ wybieramy podzbiór reprezentujący te typy pojazdu, które mogą obsłużyć daną destynację. Zbiór ten tworzy tzw. *FeasibleTypes* (FT) – dopuszczalna lista typów pojazdu. Wówczas do warunków akceptowalności trasy, związanych z ograniczeniem ładowności objętościowej i wagowej, należy doliczyć dodatkowy warunek taki, że:

$$\forall_{h \in \mathbb{R}^{(k \in \mathbb{K}_t)}} t \in FT_{i_h} \quad (6)$$



Rysunek 1. Schematyczna reprezentacja zaplanowanych tras dla problemu MultiDepot VRP z DWL [opracowanie własne]

W implementacji algorytmu można posłużyć się uproszczeniem, wprowadzając umowy poziom dopuszczalności. Zakłada się, że typy pojazdów posortowane są w kolejności od największych do najmniejszych. Przykładowo, w analizach zakładano istnienie 3 typów pojazdów oznaczanych numerami $\{1..3\}$, z których typ 1 to duży pojazd z naczepą, zaś typ 3 to mały bus towarowy. Wówczas zamiast listy FT wprowadza się wartość liczbową $Limit$ informującą, które pojazdy mogą podjechać do danej destynacji. Warunek jest spełniony, jeśli $t > Limit$.

2.2. Założenia modelu i charakterystyka zadania planistycznego

W niniejszej pracy analizowano problem marszrutyzacji polegający na zaplanowaniu tras rozwoju dóbr konsumpcyjnych (spożywczo-przemysłowych), pomiędzy 202 punkty odbioru (destynacje) zlokalizowane głównie na obszarze województwa świętokrzyskiego i ziemi radomskiej. Położenie punktów odbioru odpowiada położeniu rzeczywistych placówek handlowych, dla których znane są ograniczenia na podjazd dla pojazdów wielkogabarytowych, w szczególności ciągników siodłowych z przyczepą. Założono wariant z pojedynczym depotem – hurtownią towaru, która jednocześnie jest bazą dla pojazdów transportowych. Depot występuje na pierwszej pozycji na liście destynacji, liczącej w związku z tym łącznie 203 pozycje. Odległości pomiędzy poszczególnymi parami destynacji pozyskano z serwisu OSRM (ang. *Open Source Routing Machine*), w sposób opisany w pracach [17, 18]. Odległości reprezentowano w postaci macierzy kwadratowej asymetrycznej o wymiarze 203×203 , utożsamianej z $DistMx$. Równolegle na podstawie szacunkowych prędkości średnich na zadanym dystansie wyznaczono tzw. macierz interwałów ($TimeMx$), wykorzystywany przy szacowaniu czasu realizacji usług dostarczenia towaru. Ponadto w modelu czasowym założono dodatkowy czas związany z zajazdem do destynacji oraz czas trwania usługi wyładunku, proporcjonalny do objętości dostarczanego towaru. Zdefiniowano również slot czasowy pracy całej floty na 12 h (domyślnie w godzinach 6-18) oraz maksymalny czas trwania pojedynczej trasy (8-10 h). Ponadto, przy szacowaniu czasu trwania trasy

uwzględniono zalecane przerwy regeneracyjne dla kierowców w wymiarze 45 minut po każdych 4 h jazdy.

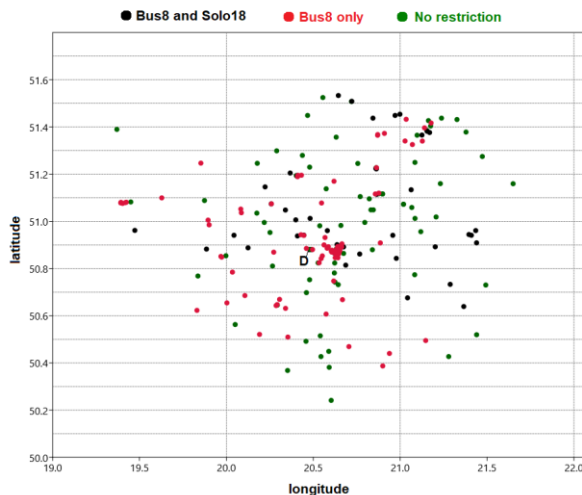
Założono, że usługi transportowe realizowane będą z udziałem trzech różnych typów pojazdów różniących się ładownością dopuszczalną wyrażoną w standardowych paletach o wymiarach 80 x 80 x 120 cm (tab. 1).

Tabela 1. Charakterystyka typów pojazdów

TYP	Nazwa	Charakterystyka	Ładowność objętościowa [palety]	Ładowność wagowa [kg]	Ilość pojazdów
1	Tir27	Ciągnik siodłowy z przyczepą	27	19950	∞
2	Solo18	Ciągnik siodłowy	18	8500	∞
3	Bus8	Bus Transporter	8	2500	∞

Źródło: opracowanie własne.

Podane parametry w dużym założeniu odpowiadają rzeczywistości potencjałowi pojazdu. Oprócz dopuszczalnej ładowności objętościowej uwzględniono opcjonalnie ładowność wagową – oba limity muszą być spełnione równocześnie, by spełnić warunki dopuszczalności rozwiązania. Jak wspomniano, na pojazdy danego typu mogą być nałożone dodatkowe warunki na dojazd (*DWL*). Założono, że jedynie pojazdy typu 3 (*Bus8*) mogą obsłużyć każdą z 202 destynacji. Pojazdy typu 1 (ciągniki siodłowe z przyczepą) mogą obsłużyć do 72 z 202 destynacji (~36%), zaś pojazdy typu 2 (ciągniki siodłowe bez przyczepy) – kolejnych 44 destynacji, tj. łącznie ok. 57% wszystkich destynacji. Pozostałe 86 destynacji (ok. 43%) mogą być wyłącznie obsłużone przez pojazdy typu 3. Poglądowe rozlokowanie destynacji przedstawiono na rysunku 2. Punkty o różnym poziomie restrykcji *DWL* oznaczono odmienną barwą. Lokalizację depotu oznaczono symbolem *D*. Z uwagi na rzeczywisty charakter problemu, rozkład destynacji charakteryzuje się pewnym poziomem klasteryzacji, której nasilenie utożsamiane jest z występowaniem ośrodków miejskich.



Rysunek 2. Szacunkowe rozlokowanie destynacji wokół depotu z zaznaczeniem poziomu ograniczeń na pojazd (*DWL*) [opracowanie własne]

Lista zamówień (wolumen i masa dostaw) została wygenerowana na podstawie reprezentatywnych danych pozyskanych z rzeczywistych danych logistycznych odnoszących się do realnych zamówień. W przeciwieństwie do klasycznie definiowanego problemu *CVRP*, łączny wolumen zamówień kierowany do pojedynczej destynacji może przekraczać ładowność pojazdów, jednocześnie ich ilość, w liczbie 423, znacznie przekracza ilość destynacji. Maksymalny rozmiar pojedynczego zamówienia wynosi 3 palety, jest zatem mniejsza od ładowności każdego z pojazdów floty. Z kolei maksymalny wolumen kierowany do pojedynczej destynacji wynosi 14, jest więc większy niż ładowność pojazdów typu 3. Oznacza to, że liczba unikalnych tzw. *route-points* może być potencjalnie większa od 202.

3. Charakterystyka algorytmu

3.1. Podstawowy wariant deterministyczny

Zgodnie z heurystyką zachłanną, analizowany algorytm bazuje na zasadzie lokalnie najkorzystniejszego doboru zlecenia. W stanie początkowym wszystkie pojazdy znajdują się w pozycji depotu, wszystkie zlecenia są w stanie *unserved*, a aktualny koszt wynosi 0. W każdym kroku podejmowana jest próba przypisania jednego zlecenia do jednego z pojazdów floty. Odbywa się to w dwóch głównych fazach. W pierwszej znajdowany jest tzw. *LocalBestHop*, czyli dopuszczalna para destynacji (*predecessor-sukcesor*), której wybór w najmniejszym stopniu zwiększa aktualną wartość kosztu, zgodnie z heurystyką „najbliższego sąsiada”. Sprowadza się to do wyboru, dla każdej kolejnej iteracji, najmniejszej wartości dystansu, będącego elementem pewnej podmacierzy zawartej macierzy *DistMx*, w której numer wiersza (*predecessor*) odnosi się do destynacji lub depotu, w której znajduje się choć jeden pojazd potencjalnie zdolny przyjąć zlecenie (pojazd aktywny), zaś numer kolumny odnosi się do destynacji, dla której istnieje jest co najmniej jedno zlecenie w stanie *unserved* (1).

$$Cost[k] = Cost[k - 1] + \min (d_{i \in HistPos, j \in ActiveDEST}) \quad (7)$$

gdzie *HistPos* oznacza aktualne pozycje pojazdów aktywnych, zaś *ActiveDEST* reprezentuje aktualny rozkład zleceń w stanie *unserved*. Założono, że w stanie początkowym wszystkie dostępne pojazdy znajdują się w depocie (destynacji nr 1). W drugiej fazie wybierana jest najkorzystniejsza para pojazd-zlecenie, preferując pojazdy i zlecenia o jak największej ładowności rozumianej jako wypadkowa ładowności objętościowej i wagowej. W szczególnym przypadku logika algorytmu sprowadza się do wyboru pary, dla której suma ładowność pojazdu i wolumen zlecenia jest największy. Dodatkowo można uwzględnić wartość skalarną reprezentującą poziom akceptowalności danego typu pojazdu przez destynację. W ogólności zaproponowano, aby najlepszą parę wybrać na podstawie maksimum wartości iloczynu 3 niezależnych czynników:

$$BestPair(t_k, n_i) = \max (\alpha_{t_k, n_i} \cdot \beta_{t_k, n_i} \cdot \gamma_{t_k}) \quad (8)$$

gdzie:

$$\alpha_{t_k, n_i} = Vmax_{t_k} + LoadVol_{n_i} \quad (9a)$$

$$\beta_{t_k, n_i} = Mmax_{t_k} + LoadMass_{n_i} \quad (9b)$$

są sumami odpowiednio ładowności i ciężaru pojazdu znajdującego się w pozycji predecesora i oraz objętości i ciężaru zamówienia kierowanego do sukcesora, natomiast ostatni z czynników wyraża się wzorem:

$$Y_{t_k} = \frac{1}{1 + \text{Limit}_{t_k}} \quad (9c)$$

gdzie Limit_{t_k} jest skalarną reprezentacją poziomu dopuszczalności danego pojazdu do bazy destynacji. Założono przy tym, że pojazdy o większych gabarytach będą z reguły charakteryzować się niższą wartością Limit , gdyż istnieje większe prawdopodobieństwo wystąpienia ograniczenia DWL na ten typ pojazdu. Taka specyfika algorytmu preferuje wykorzystanie w pierwszej kolejności pojazdów, na które prawdopodobieństwo nałożenia limitu jest największe. Tym samym mityguje to ryzyko przedwczesnego wyczerpania pojazdów o mniejszych gabarytach i wejścia w tzw. stan globalnego wstrzymania, który *de facto* jest zatrzymaniem algorytmu w stanie niekompletności.

W tym momencie należy zauważyć, że wybór najlepszej pary może być niemożliwy z kilku powodów, które można podzielić na dwie grupy:

1. przypadek „przeładowania”, związany z przekroczeniem limitu na ładowność objętościową, wagową lub dopuszczalny czas trwania trasy;
2. przypadek limitu DWL lub (opcjonalnie, gdy występuje) ograniczenia na okno czasowe dostaw.

W pierwszym przypadku uznaje, że pojazd wyczerpał swoje zasoby przewozowe (ładunek, ciężar lub czas jazdy), a realizowana przez niego trasa traktowana jest jako kompletna. W zaproponowanym rozwiązaniu dopuszcza się jednak ponowne użycie tego samego pojazdu, jeśli dostępny jest minimalny limit czasu do zakończenia slotu roboczego. Proces taki, nazwany readdycją, poprawia wykorzystanie zasobów transportowych i ułatwia agregację tras realizowanych w danym dniu roboczym przez ten sam pojazd. W związku z powyższym należy rozróżnić pojęcie liczby tras (*routes*), która zazwyczaj jest większa od liczby zaangażowanych pojazdów (*trucks*).

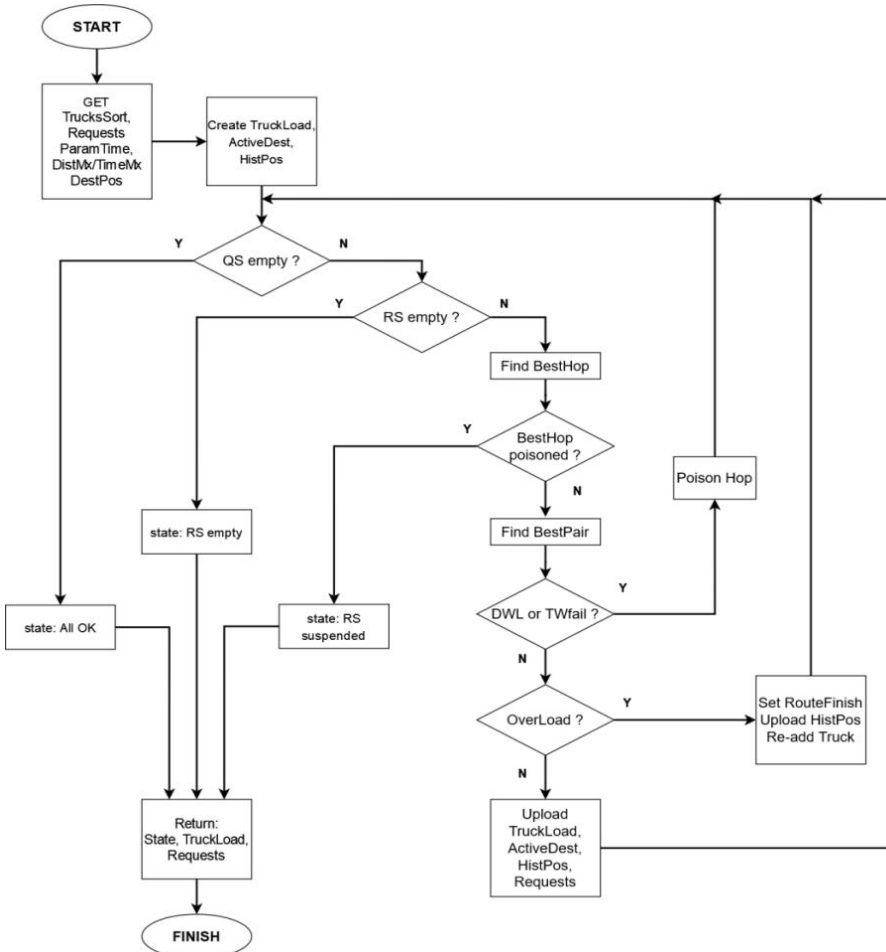
W drugim przypadku pojazd nie może zrealizować „skoku” uznanego przez algorytm za lokalnie najlepszy. W celu uniknięcia zapętlenia, zastosowano metodę „zatrucia” skoku, polegającą na zastąpieniu oryginalnej wartości dystansu wartością ponadnormatywną. W omawianej implementacji zaproponowano, by ową wartością zawsze była liczba większa o 1 od maksymalnej wartości występującej w DistMx . W rezultacie, w kolejnym kroku, algorytm wybierze inną parę predecesor-sukcesor. Równocześnie pojazd przechodzi w stan wstrzymania, jednak jest on wciąż w stanie przyjąć potencjalne zlecenia, kierowane do innych destynacji. Destrukcyjną decyzją byłoby przedwczesne zakończenie trasy dla tego pojazdu. Łatwo zauważyć, że w przypadku zatrucia skoku prowadzącego z depotu wiele pojazdów mogłoby być nieużytecznych z uwagi na podjęcie decyzji o zakończeniu trasy jeszcze przed wyjazdem w trasę.

Algorytm korzysta z danych wejściowych reprezentowanych z reguły przez tablice dwuwymiarowe, zawierające niezbędne informacje o zasobach transportowych (TrucksSort), liście zleceń (Requests), zbiorze destynacji i depotu(ów) (DestPos), tablicach interdystansów (DistMx , TimeMx) czy wreszcie zestawu współczynników dla modelu czasowego. Lista TrucksSort utożsamiana jest z tzw. przestrzenią zasobów (RS – ang. *Resource Space*), przy czym z uwagi na dopuszczalną readdycję, baza ta może być uzupełniania w trakcie trwania algorytmu. Z kolei lista Requests jest głównym

elementem tzw. przestrzeni żądań (QS – ang. *Query Space*). W trakcie wykonywania algorytmu dane zawarte w *Requests* nie są modyfikowane z wyjątkiem flagi oznaczającej obsłużenie zlecenia (*unserved/served*). Oczekiwany rezultatem jest stan, w którym wszystkie zlecenia w *Requests* przejdą w stan *served*, jednak może do tego nie dojść z dwóch głównych powodów:

1. Zasoby RS zostały wyczerpane, czego przyczyną jest zbyt mała liczba zarezerwowanych pojazdów;
2. Zasoby RS zostały wstrzymane wskutek „zatrucia” wszystkich skoków – tzw. stan globalnego wstrzymania. W przeciwieństwie do powodu pierwszego sytuacja taka jest spowodowana niewłaściwym składem floty (np. zbyt duża ilość pojazdów wielkogabarytowych w stosunku do busów) lub nieprawidłowym doбором zasobów przez algorytm (przedwczesne wyczerpanie pojazdów nisko gabarytowych).

Jak wspomniano, dzięki preferowaniu pojazdów wielkogabarytowych przy wyborze najlepszej pary zlecenie-pojazd, stan globalnego wstrzymania jest mniej prawdopodobny pod warunkiem dysponowania odpowiednimi zasobami. Ogólny schemat algorytmu przedstawiono na rysunku 3.



Rysunek 3. Schemat blokowy algorytmu Greedy [opracowanie własne]

Wynikiem zwracanym przez algorytm jest rozwiązanie w postaci zestawu tzw. *ServiceRecords (SRs)*, będącą listą unikalnych tras, reprezentowaną w postaci tablic zagnieżdżonych. Najważniejszymi elementami *SR*-a jest identyfikator trasy, identyfikator pojazdu oraz tzw. sekwencja, tj. lista zleceń ustawiona w odpowiedniej kolejności. Z uwagi na możliwość wystąpienia dwóch lub więcej zleceń do jednej destynacji, sekwencja jest szerszym pojęciem niż trasa (*route*), która zawiera jedynie listę kolejnych destynacji odwiedzanych przez pojazd. Dla większej funkcjonalności, w sekwencji przechowywane są także szacowane wartości momentów czasowych dostaw. Z kolei pojedynczy *SR* może być uzupełniony o przydatne dane statystyczne dotyczące długości trasy, przewidywalnego czasu jej trwania, sumarycznej ilości towaru, zalecany czas wyjazdu.

Powyższy algorytm zaimplementowany został w wersji demonstracyjnej w środowisku *Mathcad* z wykorzystaniem narzędzia *Programming*. Dane importowano z podręcznych plików tekstowych. Średni czas realizacji zadania optymalizacyjnego dla 423 zleceń wynosi ok. 0,25 s, przy użyciu komputera PC Intel(R) Core(TM) i7-8750H CPU @ 2,20 GHz i 8 GB RAM.

3.2. Model zaburzeń stochastycznych

Jak wspomniano, cechą charakterystyczną algorytmu zachłannego jest jego deterministyczność, w rezultacie zwraca on zawsze identyczne rozwiązanie przy identycznym stanie danych wejściowych. Z uwagi na nałożone ograniczenia czy rozkład pojazdów we flocie, algorytm zachłanny może znaleźć rozwiązania znacznie odbiegające od oczekiwanych. Dlatego też wprowadzono drobną poprawkę polegającą na użyciu losowo zmodyfikowanych wartości tablicy *DistMx*, z której algorytm korzysta przy wyznaczaniu *LocalBestHop*. Jednocześnie pozostałe parametry, a w szczególności *TimeMx* zostają zachowane, dzięki czemu znalezione rozwiązania alternatywne spełniają warunki dopuszczalności. Zaproponowaną metodykę nazwano GAMP (ang. *Greedy Algorithm with Matrix Perturbation*).

Rozważyć można kilka typów zaburzeń wartości *DistMx*, które podzielić można na multiplikatywne i addytywne, symetryczne i asymetryczne czy wreszcie pod kątem dystrybucji rozkładu (równomierne, gaussowskie i inne). Zaburzenia multiplikatywne polegają na przemnożeniu danej wartości macierzy *DistMx[i,j]* przez losowy współczynnik $c_{i,j}$ o wartości oczekiwanej równej 1. Jeśli element *DistMx[i,j]* jest mnożony przez ten sam współczynnik co *DistMx[j,i]* zaburzenie ma charakter symetryczny. Zaburzenia addytywne rozumiane są jako dodawanie wartości losowej o wartości oczekiwanej równej 0. Należy tu zaznaczyć, że zaburzeniu nie podlegają elementy diagonalne. Rozkład wartości współczynników zaburzeń może mieć charakter równomierny z zakresu $(-p/2;p/2)$, gdzie $p < 2$, lub gaussowski o zadanym odchyleniu standardowym σ . W tym drugim przypadku istnieje jednak możliwość takiego zmodyfikowania wartości dystansu, że przyjmą one może wartości ujemne. Przykład reguły generującej macierz zaburzoną w wariancie multiplikatywnym i symetrycznym, o rozkładzie równomiernym zawarto w wyrażeniu (10).

$$\begin{aligned} DistMxP[i,j] &= (1 - p/2 + rand.uni(0,p)) \cdot DistMx[i,j] \text{ dla } i > 0, j > i \\ DistMxP[j,i] &= DistMxP[i,j] \end{aligned} \quad (10)$$

gdzie $rand.uni(0,p)$ oznacza losową liczbę rzeczywistą z zakresu $0 - p$. Z kolei reguła służąca generacji macierzy zaburzonej o rozkładzie gaussowskim oparta jest o równanie (11)

$$\begin{aligned} DistMxP[i,j] &= (1 + rand.gauss(0,\sigma)) \cdot DistMx[i,j] \text{ dla } i > 0, j > i \\ DistMxP[j,i] &= DistMxP[i,j] \end{aligned} \quad (11)$$

gdzie $rand.gauss(0, \sigma)$ oznacza losową liczbę o rozkładzie gaussowskim o wartości oczekiwanej równej 0 i odchyleniu standardowym równym σ .

4. Rezultaty

4.1. Wariant deterministyczny

Z wykorzystaniem opracowanego i zaimplementowanego algorytmu dokonano przypisania 423 zleceń o łącznym wolumenie 569 palet do heterogenicznej floty transportowej złożonej z 3 typów pojazdów, scharakteryzowanych w sekcji 3. Założono, że rozmiar floty jest praktycznie nieograniczony. Listę zleceń (*Requests*) zaprezentowano w dwóch wariantach:

- z domyślnymi wartościami okien czasowych tożsamy ze slotem roboczym floty (domyślnie godz. 6-18, oznaczane *NoTW*);
- z umiarkowanymi ograniczeniami okien czasowych dla wybranych destynacji, redukującymi średni rozmiar okna z 12 do ok. 8,7 (*WithTW*).

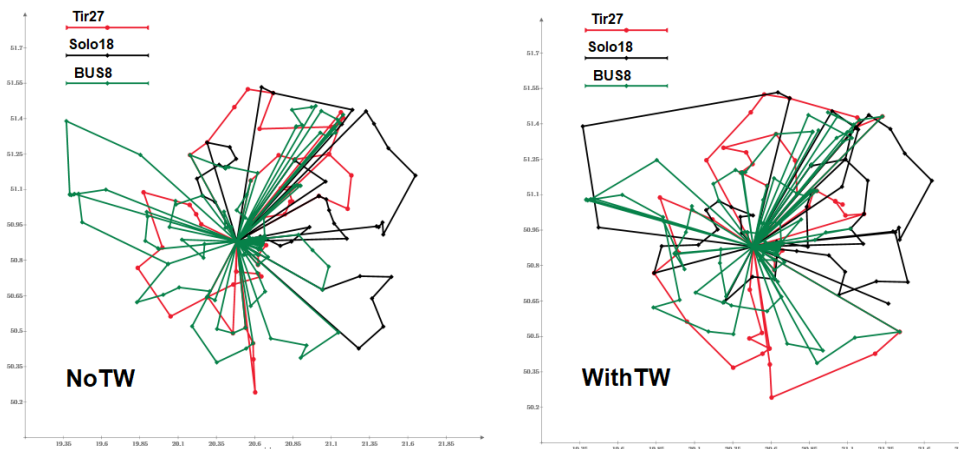
W obu przypadkach nałożono dodatkowe ograniczenie na maksymalny czas trwania pojedynczej trasy, wynoszący 10 h. W tabeli 2 przedstawiono podstawowe rezultaty dla wariantu deterministycznego.

Tabela 2. Porównanie rezultatów zwracanych przez deterministyczny algorytm zachłanny dla różnych poziomów ograniczeń i doboru flot

	TYP pojazdu	Liczba tras	Liczba wymaganych pojazdów	Liczba obsłużonych zleceń	Przewożony wolumen (palety)	Przewożony wolumen [%]	Długość trasy
DWL (NoTW)	Tir27	8	7	124	192	33,7	6418,4
	Solo18	8	6	85	122	21,4	
	Bus8	34	15	214	255	44,9	
	RAZEM	50	28	423	569	100	
DWL (WithTW)	Tir27	8	7	103	160	28,2	7621,7
	Solo18	11	9	110	159	27,9	
	Bus8	36	17	210	250	43,9	
	RAZEM	55	33	423	569	100	
No DWL NoTW	Tir27	23	12	423	569	100	3886,0
	Solo18	0	0	0	0	0	
	Bus8	0	0	0	0	0	
BUS only	Bus8	74	20	423	569	100	8819,9

Źródło: opracowanie własne.

Dla celów porównawczych przeanalizowano hipotetyczną sytuację, w której wszystkie ograniczenia *DWL* zostają zniesione. W takim przypadku, zgodnie z oczekiwaniami, algorytm faworyzuje wykorzystanie pojazdów wielkogabarytowych, które praktycznie przejmują 100% zadań transportowych. Pozwala to na dość istotne zredukowanie całkowitej długości trasy o ok. 40%. W przeciwstawnej skrajnej sytuacji, w której wykorzystuje się wyłącznie pojazdy małogabarytowe, długość trasy wzrasta o blisko 40%. Poglądowy przebieg tras przedstawiono na rysunku 4. Trasy realizowane przez pojazdy tego samego typu oznaczono jednakowym kolorem.



Rysunek 4. Poglądowa wizualizacja tras realizowanych przez pojazdy poszczególnych typów [opracowanie własne]

4.2. Wariant stochastyczny z zaburzeniem o rozkładzie równomiernym

W przypadku wprowadzanych zaburzeń multiplikatywnych o rozkładzie równomiernym (formuła 10), przeprowadzono po $N = 100$ prób losowych dla każdej wartości parametru $p = 0,001; 0,003; 0,01; 0,03; 0,1; 0,3$ oraz 1. Łącznie wykonano więc 700 wywołań algorytmu zachłannego, a każdy wynik w postaci kosztu kilometrażu, liczby unikalnych tras oraz liczby niezbędnych pojazdów odniesiono do rezultatu oryginalnego prezentowanego w tabeli 2. Analizy przeprowadzono dla dwóch wariantów ograniczeń okien czasowych, prezentowanych w tejże tabeli.

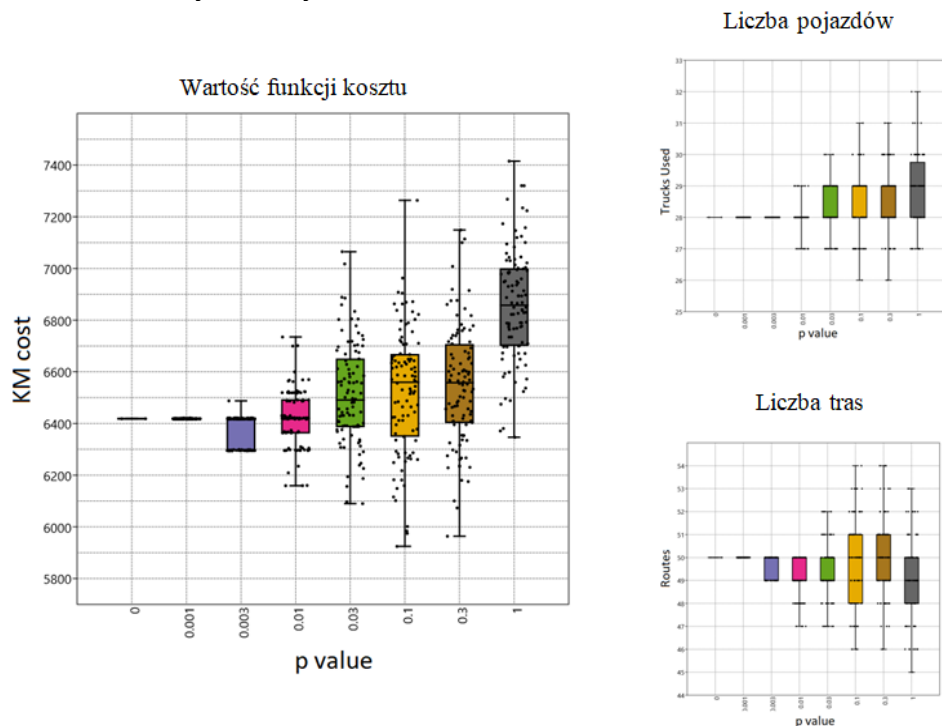
W tabeli 3 przedstawiono zestawienie wybranych parametrów statystycznych, pogrupowanych względem wartości p , dla uzyskanej wartości funkcji kosztu. Globalnie najkorzystniejszy wynik dla wariantu *NoTW* uzyskano dla przypadku $p = 0,1$, gdzie był on o ok. 7,7% niższy od rozwiązania oryginalnego. Natomiast dla wariantu *WithTW*, najkorzystniejszy przypadek przypadł dla wartości dla $p = 0,3$, zaś sam wynik był aż ponad 14% niższy. Generalnie wzrost p prowadzi do zwiększenia wariancji wyników, co przekłada się na zwiększenie szansy na znalezienie korzystniejszych rezultatów. Jednak przy wartościach p bliskich 1 zakres wartości przesuwa się w górę, dlatego ta wartość nie jest rekomendowana do dalszego użycia w zaproponowanej procedurze. Jak należało o się spodziewać, zbyt małe wartości p (rzędu 0,001) mogą pomijalnie wpływać na wygenerowanie rozwiązań alternatywnych.

Tabela 3. Wyniki statystyczne dla wariantu stochastycznego z zaburzeniem o rozkładzie równomiernym, z podziałem na wartość amplitudy zaburzenia p

	p	0	0,001	0,003	0,01	0,03	0,1	0,3	1
		(oryg)							
	N	1	100	100	100	100	100	100	100
NoTW	Min	6418,4	6418,4	6294,2	6159,4	6090,0	5924,5	5963,8	6346,8
	Max	6418,4	6421,5	6487,8	6734,7	7065,0	7263,5	7148,8	7415,2
	Mean	6418,4	6418,5	6385,2	6421,7	6513,7	6523,6	6550,2	6857,0
	SD	N/A	0,4	56,3	107,9	193,2	242,9	227,2	215,5
WithTW	Min	7621,7	7405,4	7070,5	6809,2	6609,4	6585,7	6535,9	6834,5
	Max	7621,7	7622,0	7625,5	7738,3	7871,7	8124,3	8086,5	8766,0
	Mean	7621,7	7606,6	7507,6	7320,2	7157,7	7213,6	7187,0	7588,9
	SD	N/A	55,5	148,0	196,0	241,7	296,9	310,5	324,8

Źródło: opracowanie własne.

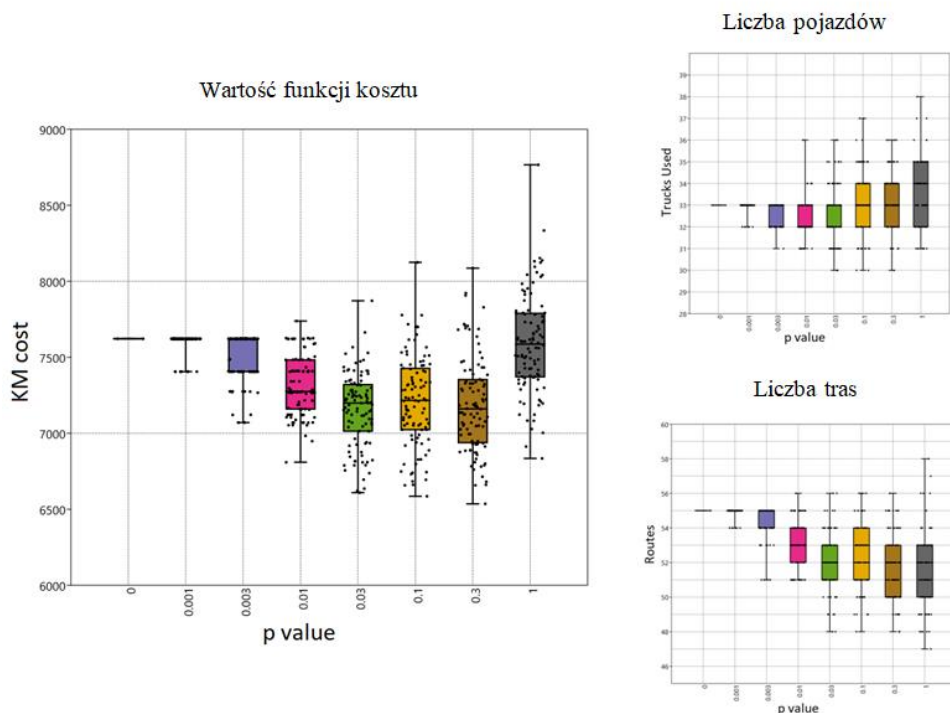
Na rysunku 5 przedstawiono zobrazowanie uzyskanych rezultatów dla wariantu *NoTW*, w postaci wykresów *Box&Jitter*, z podziałem na wartość amplitudy zaburzenia (p). Obrazowanie i analizy wykonano za pomocą darmowego oprogramowania PAST [19]. Zaznaczono wartości mediany, odstęp międzykwartylowy oraz wartości minimalnych i maksymalnych dla wartości funkcji kosztu kilometrażowego, liczby wymaganych pojazdów oraz liczby unikalnych tras.



Rysunek 5. Reprezentacja wyników symulacji w postaci wykresów *Box&Jitter* z podziałem na wartość amplitudy zaburzenia – wariant *NoTW*, zaburzenie o rozkładzie równomiernym [opracowanie własne]

Wykres odzwierciedla wcześniejsze obserwacje o pomijaniu małym wpływie zaburzeń o małych amplitudach oraz niekorzystne zwiększenie zakresu uzyskiwanych wyników dla dużych amplitud zaburzeń. Jak zaprezentowano na wykresie, znaczna część prób zaburzeń skutkowałą uzyskaniem wyniku korzystniejszego zarówno pod względem kosztu końcowego, liczby użytych pojazdów czy liczby niezależnych tras. Jednakże wartości mediany, stanowiącej statystyczne uśrednienie wyniku lokuje się zazwyczaj powyżej poziomu odniesienia. W przypadku liczby tras zaobserwowano ich wyraźną redukcję dla skrajnej wartości zaburzenia ($p = 1$), jednakże otrzymane wyniki nie przekładają się na koszt kilometrażowy. Należy zatem podkreślić, że liczba tras nie powinna być wyznacznikiem efektywności danej próby pod kątem minimalizacji funkcji celu.

Jeszcze korzystniejszy wpływ procedury zaburzenia można zaobserwować dla wariantu *WithTW*. Wartość oryginalna funkcji kosztu w tym przypadku jest aż o ok. 18,7% wyższa od wyniku oryginalnego dla wariantu *NoTW*. Na przykładzie analogicznego wykresu rozrzutu dla wariantu *WithWT* można założyć, że rozwiązanie oryginalne zdaje się wypadać wyjątkowo niekorzystnie na tle prób stochastycznych – rysunek 6. Dla wartości p z zakresu 0,01-0,3 ponad 90% prób kończyło się wynikiem korzystniejszym od wyniku oryginalnego.



Rysunek 6. Reprezentacja wyników symulacji w postaci wykresów *Box&Jitter* z podziałem na wartość amplitudy zaburzenia – wariant *WithTW*, zaburzenie o rozkładzie równomiernym [opracowanie własne]

W tabeli 4 przedstawiono podsumowanie najważniejszych wyników uzyskanych podczas prób zaburzenia macierzy interdystansów. Porównano wartości rozwiązania oryginalnego, z wartością najlepszego znalezionego rozwiązania oraz ze średnią

arytmetyczną rezultatów dla podgrupy 400 przypadków odpowiadającym wartościom amplitudy zaburzeń $p = 0,01; 0,03; 0,1$ i $0,3$. Zgodnie z wcześniejszymi obserwacjami pominięto wartości najmniejsze i największe tej amplitudy. W tabeli przedstawiono parametry rozwiązania oryginalnego, rozwiązania najlepsze, średnia z wywołań, a także odsetek wywołań, które zwracały rezultat niższy (lepszy) od oryginalnego.

Jak wspomniano, najlepsze znalezione rozwiązanie dla wariantu *NoTW* charakteryzowało się wartością kosztu niższą o ok. 7,7% od rozwiązania oryginalnego (5924,5 vs. 6418,4), natomiast średnia wartość dla wybranej grupy była jednak nieco większa (6502,2). Około 1/3 prób kończyła się wynikiem korzystniejszym od rozwiązania oryginalnego. Próby takie nazwano „sukcesami”. W blisko połowie przypadków uzyskano niższy udział tras, jednak jak zauważono powyżej, nie jest to wyznacznik stopnia optymalizacji.

Tabela 4. Wyniki statystyczne dla wariantu stochastycznego z zaburzeniem o rozkładzie równomiernym

Wariant	Parametr	Rozwiązanie oryginalne	Najlepsze znalezione rozwiązanie	Średnia dla $p \in (0,01-0,3)$	Odsetek „sukcesów”
DWL (NoTW)	Pojazdy	28	26	28,3 (+0,9%)	14,8
	Trasy	50	45	49,6 (-0,8%)	47,5
	Koszt KM	6418,4	5924,5 (-7,7%)	6502,2 (+1,3%)	33,5
DWL (WithTW)	Pojazdy	55	47	52,3 (-4,9%)	43,8
	Trasy	33	30	32,8 (-0,8%)	87,8
	Koszt KM	7621,7	6535,9 (-14,2%)	7219,45 (-4,5%)	92

Źródło: opracowanie własne.

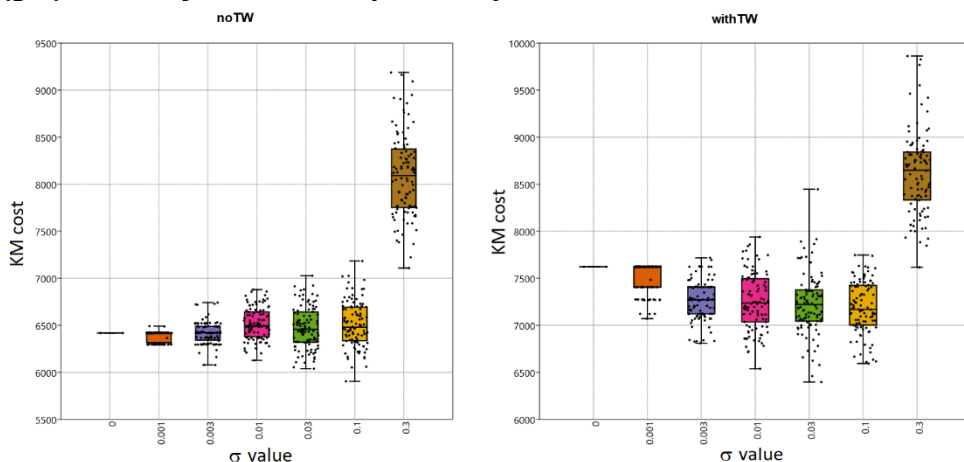
Warta uwagi jest jednak redukcja liczby wymaganych pojazdów, gdyż ich udział dodatkowo wpływa na sumaryczny koszt floty, który jednak nie był uwzględniony w zaproponowanej funkcji kosztu kilometrażowego. W przypadku wariantu *WithTW* najlepsze rozwiązanie charakteryzuje się kosztem o ok. 14,2% niższym, zaś ponad 92% prób kończyło się uzyskaniem rezultatu korzystniejszego niż oryginalny. Średnia wartość zwracanego kosztu również była niższa od wartości początkowej o ok. 4,5%. Spowodowane jest to prawdopodobnie tzw. stopniem korzystności deterministycznego rozwiązania zachłannego na tle przestrzeni rozwiązań dopuszczalnych. Średnie rozwiązanie stochastyczne dla wariantu *WithTW* jest już tylko o ok. 11% większe od analogicznej wartości dla wariantu *NoTW*.

4.3. Wariant stochastyczny z zaburzeniem o rozkładzie gaussowskim

Analogiczne analizy przeprowadzono dla zaburzeń o rozkładzie gaussowskim. Przeprowadzono po $N = 100$ wywołań algorytmu *Greedy* z zaburzeniami opisywanymi zależnością (10). Zgodnie ze specyfiką rozkładu normalnego, można założyć, że wartość σ powinna być ok. 3-krotnie mniejsza od szacunkowej amplitudy zaburzeń. W związku z tym symulacje przeprowadzono dla wartości $\sigma = 0,0005; 0,0015; 0,005; 0,015; 0,05; 0,015$, co daje 600 wywołań.

Na rysunku 7 przedstawiono reprezentację populacji wyników kosztu KM dla zaburzenia gaussowskiego. Zestawiono tu oba warianty *NoTW* i *WithTW*, dla których profil jest generalnie dość zbliżony do analogicznych rezultatów uzyskanych dla symulacji

zaburzeń o rozkładzie równomiernym. Wyjątkiem jest przypadek największej wartości parametru σ , dla którego wyniki wydają się negatywnie odstępować od pozostałej reprezentacji. Przyczyną jest duże prawdopodobieństwo osiągnięcia wartości ujemnych w macierzy interdystansów dla tej wartości zaburzenia. W efekcie wartości ujemne okazały się bardziej konkurencyjne w procedurze wyszukiwania najlepszych skoków od wartości diagonalnych. W efekcie algorytm zatracił naturalną cechę agregacji zleceń kierowanych do tej samej destynacji. Dlatego zbyt duża wartość zaburzenia nie jest rekomendowana lub też należy wprowadzić dodatkowe ograniczenie uniemożliwiające osiągnięcie zerowej wartości interdystansu w procesie zaburzenia.



Rysunek 7. Reprezentacja wyników symulacji w postaci wykresów *Box&Jitter* z podziałem na wartość amplitudy zaburzenia – warianty *NoTW* vs. *WithTW*, zaburzenie o rozkładzie gaussowskim [opracowanie własne]

Dla obu typów zaburzeń dokonano porównania najważniejszych parametrów rozkładu statystycznego otrzymanych wyników. W obu przypadkach ograniczono się do 4 najkorzystniejszych serii (łącznie 400 wywołań), odrzucono natomiast wartości skrajne. Wyniki przedstawiono w tabeli 4.

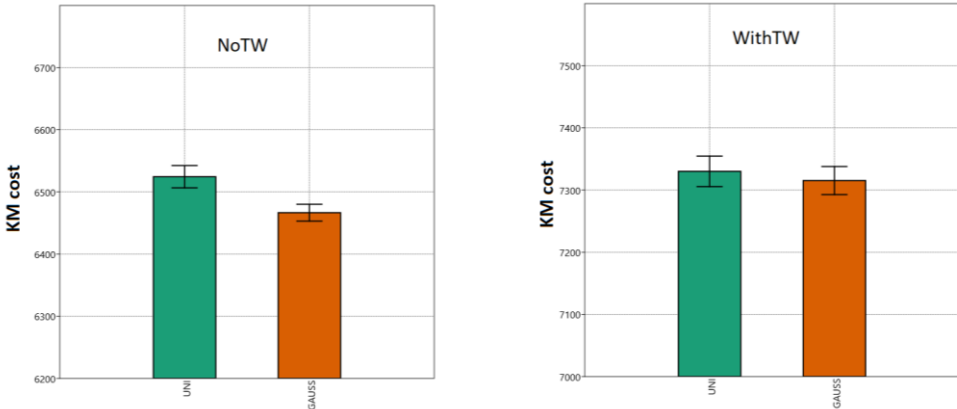
Tabela 4. Porównanie statystycznych wyników dla dwóch wersji zaburzeń

Wariant	Parametry kosztu	Zaburzenia UNI	Zaburzenia GAUSS
DWL (No TW)	Minimum	5924,5	5799,6
	Średnia	6524,5	6466,6
	Odch.std	224,8	170,5
	Odsetek „sukcesów”	29,3%	32%
DWL (With TW)	Minimum	6535,9	6395,8
	Średnia	7330,1	7315,3
	Odch.std	307,6	282,8
	Odsetek „sukcesów”	78,2%	75%

Źródło: opracowanie własne.

Stwierdzono, że średni uzyskany wynik jest niższy (korzystniejszy) dla zaburzeń typu gaussowskiego, co pozwala rekomendować użycie tego zaburzenia jako bardziej skutecznego. Różnice są mniejsze dla wariantu *WithTW*. Kolejną przewagą dla zaburzeń typu gaussowskiego wydaje się być zdolność do znalezienia najniższej wartości dla rozwiązania w zbiorze wywołań.

Wyniki tabeli 4 zaprezentowano także w postaci wykresu z zaznaczonym zakresem odpowiadającym 95% przedziałowi ufności (95%CI) – rysunek 8. W wariancie *NoTW* zakres uzyskanych wyników wydaje się być statystycznie niższy dla przypadku zaburzeń o rozkładzie gaussowskim, natomiast dla wariantu *WithTW*, przedziały ufności wyraźnie się zazębiają, nie pozwalając na jednoznaczne potwierdzenie hipotezy o przewadze rozkładu gaussowskiego. Potwierdzają to przeprowadzone testy statystyczne (*t-Studenta*, *Tukey'a*). Dla wariantu *NoTW* uzyskano wartość $t \sim 5$ i $p(\text{same})$ na poziomie 10^{-7} , co pozwala obronić hipotezę o przewadze zaburzenia gaussowskiego. Analogicznie dla wariantu *WithTW*, wartość to jedynie ~ 0.87 , a prawdopodobieństwo, że próby pochodzą z tej samej populacji oszacowano na ok. 37%, tj. znacznie więcej niż standardowo stosowany limit 5%.



Rysunek 8. Porównanie statystycznych wyników dla dwóch wersji zaburzeń [opracowanie własne]

Pomimo niejednoznaczności wyników, zaburzenia o charakterze gaussowskim zdają się być lepszym kandydatem do zastosowań w analizowanej metodzie. Zaleca się jednak ograniczenie amplitudy zaburzeń, reprezentowanych przez odchylenie standardowe σ do wartości nie większej niż 0,1 oraz zastosowanie dodatkowego ograniczenia uniemożliwiającego osiągnięcie wartości ujemnych, jako zjawiska wyraźnie negatywnie wpływającego na wyniki.

5. Podsumowanie

W rozdziale przedstawiono wpływ zaburzeń wartości interdystansów na zachowanie się deterministycznego algorytmu zachłannego, stosowanego do rozwiązania rzeczywistego problemu marszrutyzacji dla zróżnicowanej floty transportowej z ograniczenia na podjazd i okna czasowe dostaw. Algorytm zachłanny pozwala na szybkie uzyskanie wyniku przybliżonego, który zazwyczaj jednak znacząco odbiega od rozwiązania optymalnego. Jak pokazano, losowe zaburzenie oryginalnych (rzeczywistych) wartości odległości między obiektami docelowymi pozwala uzyskać rozwiązania alternatywne,

spełniające warunek dopuszczalności, dla których umowny koszt trasy (tutaj tradycyjnie stosowany kilometr) może być często niższy od rozwiązania początkowego. Analizowano multiplikatywne zaburzenia losowe o rozkładzie równomiernym i gaussowskim, dla których określano najkorzystniejsze wartości amplitudy tych zaburzeń. Generalnie można rekomendować dość dużą rozpiętość amplitud zaburzeń, przy czym nie powinny one prowadzić do uzyskania wartości ujemnych dystansów, które zdiagnozowano jako wyraźnie niekorzystne. Na bazie 400 prób dla każdego z typów zaburzeń, stwierdzono, że zaburzenia gaussowskie wydają się korzystniejsze i mogą być rekomendowane do zastosowania w analizowanej metodzie. Jak pokazano, pozwalają na znalezienie rozwiązania, którego koszt jest nawet o kilkanaście % niższy od rozwiązania oryginalnego. Dodatkowo stwierdzono, że algorytm zachłanny w zależności od parametrów stosowanej instancji problemu może prowadzić do rozwiązań o różnym tzw. poziomie korzystności. Jeśli rozwiązanie oryginalne cechuje się dużym poziomem korzystności, wprowadzone zaburzenia w większości prowadzą do rozwiązań gorszych. Jednak w przypadku niskiego poziomu korzystności, rekomenduje się kilkukrotne wywołanie algorytmu zachłannego, co powinno skutkować uzyskaniem co najmniej jednego wyniku o niższym koszcie. Z uwagi na dość szybkie działanie algorytmów zachłannych, w praktyce istnieje dość duży potencjał czasowy, pozwalający na wielokrotne wywołanie algorytmu w warunkach zaburzenia macierzy dystansów. Zaproponowana metodyka GAMP może zatem służyć jako zmodyfikowany sposób rozwiązywania problemów kombinatorycznych z rodziny VRP. Z drugiej strony, proponowaną procedurę można potraktować jako metodę generowania populacji startowej dla bardziej zaawansowanych algorytmów genetycznych czy rojowych.

Podziękowania

Prezentowane badania zostały zrealizowane w ramach projektu POIR.04.01.04-00-0091/19 pt. „Target – multifunkcyjny system dla flot transportowych”, finansowane ze środków Programu Operacyjnego Innowacyjny Rozwój, poddziałanie 4.1.4 (Projekty Aplikacyjne).

Literatura

1. Hota J., Debjani G.D., *Workforce Analytics Approach: An Emerging Trend of Workforce Management*, AIMS International Journal of Management, 7(3), 2013, s. 167-179.
2. Shehab E.M., Sharp M.W., Supramaniam L., Spedding T.A., *Enterprise resource planning: An integrative review*, Business Process Management Journal, 10(4), 2004, s. 359-386.
3. Zhang H., Ge H., Yang J., Tong Y., *Review of Vehicle Routing Problems: Models, Classification and Solving Algorithms*, Archives of Computational Methods in Engineering, 29, 2022, s. 195-221.
4. Konstantakopoulos G.D., Gayialis S.P., Kechagias E.P., *Vehicle routing problem and related algorithms for logistics distribution: a literature review and classification*, Operational Research, 22, 2020, s. 2033-2062.
5. Legrain A., Bouarab H., Lahrichi N., *The Nurse Scheduling Problem in Real-Life*, Journal of Medical System, 39(1), 2015, s. 160.
6. Yu Y., *A Research Review on Job Shop Scheduling Problem*, E3S Web of Conferences, 253, 2021, s. 2024.
7. Xiong H., Shi S., Ren D., Hu J., *A survey of job shop scheduling problem: The types and models*, Computers & Operations Research, 142, 2022, s. 105731.

8. Kellerer H., Pferschy U., Pisinger D., *Knapsack Problems*, Springer, Berlin Heidelberg New York 2004, ISBN 3-540-40286-1.
9. Mittal K., Jain A., Vaisla K.S., *Understanding Reviewer Assignment Problem and its Issues and Challenges*, 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), 2019, s. 1-6.
10. Held M., Karp R.M., *A dynamic programming approach to sequencing problems*, Journal for the Society for Industrial and Applied Mathematics, 1(10), 1962, s. 196-210.
11. Cormen T.H., Leiserson C.E., Rivest R.L., Stein C., *Introduction to Algorithms*, The MIT Press Cambridge, Wydanie 3, Massachusetts London, England 2009, ISBN 978-0-262-03384-8.
12. Kulkarni S., Sohani N., Sehta N., *Capacitated Vehicle Routing Using Nearest Neighbor Algorithm in Supply Chain*, International Journal of Engineering Research & Technology (IJERT), 3(5), 2014, s. 1331-4.
13. Sarwono A.A., Ai T.J., Wigat S.S., *Combination of nearest neighbor and heuristics algorithms for sequential two dimensional loading capacitated vehicle routing problem IOP Conference Series: Materials Science and Engineering*, 166, 2017, s. 12029.
14. Wijharnasir A.P.A., Alghasi M.A., Pujawan N., *Greedy approach for solving the capacitated vehicle routing problem of liquefied natural gas distribution to power plants*, MATEC Web of Conferences – ICFMCE, 272, 2018.
15. Godart A., Manier H., Bloch C., Manier M.A., *A greedy based algorithm for a bi-objective Pickup and Delivery Problem with Transfer*, International Conference on Systems, Man and Cybernetics, Oct 2019, Bari, Italy, hal-03221895.
16. Masudin I., Sa'diyah R.F., Utama D.M., Restuputri D.P., Jie F., *Capacitated Vehicle Routing Problems: Nearest Neighbour vs. Tabu Search*, International Journal of Computer Theory and Engineering, 11(4), 2019, s. 76-79.
17. Belka R., Godlewski M., *Vehicle Routing Optimization System with Smart Geopositioning Updates*, Applied Science, 11, 2021, s. 10933.
18. Belka R., Godlewski M., *Distance Matrix Generation for Dynamic Vehicle Routing Optimization in Transport Fleets Management*, [w:] Pandit M., Gaur M.K., Rana P.S., Tiwari A. (red.), *Artificial Intelligence and Sustainable Computing. Algorithms for Intelligent Systems*, Springer, Singapore 2022.
19. Hammer Ø., Harper D.A.T., Ryan P.D., *PAST: Paleontological statistics software package for education and data analysis*, Palaeontologia Electronica, 4(1), 2001.

Niedeterministyczny algorytm zachłanny do zastosowań w strategii marszrutyzacji – studium przypadku

Streszczenie

Szybkie i skuteczne planowanie tras dla flot transportowych jest jednym z ważniejszych problemów logistycznych. Odpowiednie przypisanie zadań dostaw/odbioru towaru do właściwych pojazdów skutkuje korzyścią w postaci redukcji bezpośrednich i pośrednich kosztu usług transportowych. Zagadnienie, znane powszechnie jako problem marszrutyzacji (*Vehicle Routing Problem* – VRP) jest jednym z np-trudnych problemów optymalizacji kombinatorycznej, do rozwiązań którego stosuje się są liczne metody heurystyczne (klasteryzacja, przeszukiwanie lokalne itp.) czy i metaheurystyczne (ewolucyjne, genetyczne, rojowe i inne). Jedną z prostszych i dość popularnych, ze względu na swoją uniwersalność, metod optymalizacji jest tzw. algorytm „zachłanny”, charakteryzujący się dużą szybkością w poszukiwaniu wstępnych rozwiązań problemu marszrutyzacji. W pracy przeprowadzono badania nad implementacją algorytmu zachłannego do rozwiązania rzeczywistego problemu dostaw 569 jednostek towaru do ponad 200 punktów dostaw z użyciem heterogenicznej, tj. zróżnicowanej pod względem ładowności floty w ilości ok. 30 pojazdów. Do standardowych ograniczeń związanych z ładownością nałożono dodatkowe, dotyczące okien czasowych oraz limitu na podjazd. Klasyczny wariant deterministyczny algorytmu zmodyfikowano, wprowadzając stochastyczne zaburzenia wartości dystansów między-punktowych. W efekcie algorytm odnajdywał alternatywne rozwiązania, z których znaczna część charakteryzowała się korzystniejszym wynikiem w stosunku

do rozwiązania podstawowego. Przeprowadzono badania wpływu poziomu i typu zaburzeń na dystrybucję otrzymanych wyników. Postuluje się, że metoda może posłużyć zarówno do stosunkowo szybkiego znajdowania rozwiązań przybliżonych, jak również do generowania populacji startowych dla bardziej zaawansowanych algorytmów ewolucyjnych.

Słowa kluczowe: marszrutyzacja, optymalizacja, algorytm zachłanny, logistyka

Non-deterministic greedy algorithm for applications in routing strategy – a case study

Abstract

Fast and effective route planning for transport fleets is one of the most important logistics problems. Appropriate assignment of delivery/collection requests to the right vehicles results in the benefit of reducing the direct and indirect cost of transport services. The problem, commonly known as the Vehicle Routing Problem (VRP), is one of the NP-hard problems of combinatorial optimization, to which numerous heuristic methods (clustering, local search, etc.) or metaheuristic (evolutionary, genetic, swarm etc.) methods are used. One of the simpler and quite popular, due to its universality, optimization methods is the so-called "greedy" algorithm, characterized by high speed in the search for initial solutions to the VRP problem. In the work, research was carried out on the implementation of a greedy algorithm to solve the real problem of delivering 569 units of goods to over 200 delivery points using a heterogeneous, i.e. diversified in terms of capacity, fleet of about 30 vehicles. In addition to the standard limitations related to the load capacity, additional restrictions regarding time windows and the limit for the driveway have been imposed. The classic deterministic variant of the algorithm was modified by introducing stochastic perturbations of the distances between points. As a result, the algorithm found alternative solutions, where many of which were characterized by a more favourable result compared to the basic solution. Studies on the impact of the level and type of disturbances on the distribution of the obtained results were carried out. It is postulated that the method can be used both to find approximate solutions relatively quickly, as well as to generate starting populations for more advanced evolutionary algorithms.

Keywords: routing, optimization, greedy algorithm, logistics

Głosowe interfejsy użytkownika: zastosowania, projektowanie, zagrożenia

1. Wprowadzenie

Dynamiczny wzrost wydajności sprzętu komputerowego spowodował, że w ostatnich latach rozpoznawanie i generowanie mowy naturalnej stało się dostępne zarówno dla komputerów biurowych (laptopów czy stacjonarnych), jak i popularnych urządzeń mobilnych. Syntetyzowanie mowy naturalnej osiągnęło już jakość, dzięki której rezultat jest trudny do odróżnienia od nagrania ludzkiego lektora. Natomiast rozpoznawanie mowy może być prowadzone z dokładnością pozwalającą na efektywne dyktowanie tekstu. W związku z powyższym w coraz szerszym zakresie obserwujemy próby wdrażania interfejsów użytkownika opartych na głosie zamiast standardowych urządzeń wejścia/wyjścia, takich jak monitor, klawiatura, myszka czy ekran dotykowy. Rozwiązania interakcji człowiek-komputer (ang. HCI, *Human-Computer Interaction*) z wykorzystaniem głosu naturalnego można spotkać w sklepowych kasach samoobsługowych, samochodach, inteligentnych głośnikach czy popularnych urządzeniach mobilnych. W większości przypadków opierają się one na odpowiednim słowniku poleceń. Generowane informacje wynikowe są natomiast prezentowane w postaci graficznej, słownej lub mieszanej w zależności od kontekstu. Pewne niedoskonałości takiego podejścia wydają się obecnie trudne do zniwelowania, nie ma bowiem możliwości zapewnienia całkowitej izolacji operatora i zmienne dźwięki otoczenia zawsze mogą powodować zakłócenia w takiej komunikacji. Jednak wszędzie tam, gdzie można zagwarantować użytkownikowi systemu sterowanego głosem odpowiedni poziom prywatności, można też osiągnąć rezultaty przewyższające standardowe pośrednictwa użytkowe. Zaprojektowana w ten sposób interakcja gwarantuje doświadczenie użytkownika (ang. UX, *User Experience*) zbliżone do konwersacji międzyludzkiej, redukując w znacznym stopniu konieczność szkolenia i zapoznania się z zasadami działania systemów informatycznych. Jest to szczególnie wyraźnie widoczne w rozwiązaniach wspierających osoby niepełnosprawne ruchowo lub wzrokowo, dla których alternatywne metody komunikacji z systemem komputerowym są często jedynym akceptowalnym rozwiązaniem. Przykładem takiego interfejsu może być środowisko edukacyjne wspierające dyktowanie i edycję złożonych dokumentów matematycznych, stanowiące efektywne narzędzie komunikacji człowiek-komputer z wykorzystaniem mowy naturalnej.

2. Krótki przegląd zastosowań głosowych interfejsów użytkownika

Głosowe interfejsy użytkownika znajdują szerokie zastosowanie we współczesnej technologii i urządzeniach codziennego użytku. Syntetyczna mowa wykorzystywana jest do odczytywania komunikatów, poleceń oraz dłuższych tekstów, stwarzając przy

¹ agnieszka.bier@polsl.pl, Katedra Matematyki, Wydział Matematyki Stosowanej, Politechnika Śląska, www.polsl.pl.

² zdzislaw.sroczyński@polsl.pl, Katedra Zastosowań Matematyki i Metod Sztucznej Inteligencji, Wydział Matematyki Stosowanej, Politechnika Śląska, www.polsl.pl.

tym złudzenie komunikacji z żywym lektorem. W ostatnich latach obserwuje się masowe użycie botów głosowych dla celów marketingowych, w tym automatyzacji sprzedaży, propagowania ofert, a także obsługi i nawigacji w automatycznych centrach obsługi klienta (ang. *call center*), poczcie głosowej czy bankowych systemach transakcyjnych. Interfejs głosowy jest także istotnym elementem automatycznych tłumaczy, które mogą zarówno tłumaczyć tekst mówiony w czasie rzeczywistym, jak i syntezywać poprawną wymowę fragmentów tłumaczonych jednocześnie na wiele różnych języków.

Głosowe komunikaty są nieodzownym elementem interaktywnych samoobsługowych terminali POS (ang. *Point of Sale*), czyli kas samoobsługowych. Prowadzą one klienta przez proces rejestracji produktów w koszyku, płatności oraz rozwiązywania pojawiających się problemów. Z kolei komunikacja człowiek-maszyna za pomocą interfejsu głosowego, tj. dyktowania poleceń, wiadomości tekstowych czy komunikatów sterujących stała się powszechnym i wygodnym rozwiązaniem, zwłaszcza w okolicznościach oraz na urządzeniach, które ze swojej natury uniemożliwiają łatwe wprowadzenie tekstu za pomocą klawiatury. Należy tu wymienić na przykład komunikaty sterujące dla urządzeń pokładowych w nowoczesnych samochodach czy inteligentnych urządzeniach multimedialnych (telewizorów, głośników, telefonów komórkowych) [1]. Niejednokrotnie komunikacja głosowa jest dwukierunkowa: możliwe jest zarówno podyktowanie wiadomości e-mail lub SMS, jak również jej syntetyczne odczytanie przez urządzenie. Przedstawione rozwiązania są także wykorzystywane dla wsparcia osób niepełnosprawnych: niedowidzących i niedosłyszących czy też z niedowładami kończyn górnych [2]. W powszechnym użytku są systemy, takie jak Siri czy Cortana oraz opcje ułatwienia dostępności i czytniki ekranu: „VoiceOver” w systemie Apple iOS, „Talkback” w systemie Google Android [3], „Dyktowanie i sterowanie urządzeniem” w systemie Windows, a także audiodeskrypcja i/lub transkrypcja plików graficznych i wideo osadzonych na stronach WWW.

Ze względu na tryb pracy, systemy wyposażone w interfejsy głosowe możemy podzielić na nasłuchujące ciągle oraz nasłuchujące na żądanie. W przypadku tych pierwszych interfejs głosowy jest podstawowym kanałem komunikacji z urządzeniem (np. w inteligentnych głośnikach), a konwersację rozpoczyna pewna charakterystyczna, umowna fraza, np. „Hej, Siri”. Z kolei w drugiej grupie urządzeń nasłuch komend włączany jest po wykonaniu pewnej dodatkowej czynności, jak na przykład uruchomienie trybu nasłuchu w smartfonach lub aplikacjach samochodowych, a trwająca konwersacja sygnalizowana jest odpowiednią ikoną, najczęściej odwołującą się do wzorca mikrofonu lub „mówiącej głowy”.

3. Informacja zwrotna

W porównaniu z rozpoznawaniem treści mowy z nagrania dźwiękowego, wykorzystanie głosu jako kanału zwrotnego w komunikacji człowiek-komputer jest łatwiejsze algorytmicznie, mniej wymagające technicznie i z tego względu było w praktyce implementowane jako pierwsze zastosowanie interfejsów werbalnych. W najprostszej postaci system może zawierać nagrania audio całych fraz lub pojedynczych słów przedyktowanych przez lektora i składać z nich odpowiednie wiadomości. Współcześnie wykorzystuje się częściej mowę całkowicie syntetyczną, a systemy wypowiadające w ten sposób tekst w języku naturalnym (ang. *text-to-speech*, TTS) dorównują jakością

żywym lektorom. Poza oczywistą korzyścią z wdrożenia takiej formy komunikacji dla osób niepełnosprawnych wzrokowo, istotne wydają się też aspekty psychologiczne, zmniejszające barierę niedostępności i nieufności, szczególnie w przypadku mało doświadczonych użytkowników. Można stosunkowo niewielkim kosztem przygotować komunikację w wielu językach narodowych, nawet dla osób, którym przeczytanie dłuższego tekstu sprawia nieproporcjonalny wysiłek, co ma miejsce przykładowo w systemach e-learningu przeznaczonych dla najmłodszych.

Systemom przedstawiającym werbalne komunikaty zwrotne najczęściej towarzyszą elementy graficzne, czy to w postaci tekstu, czy też ikonografiki. Czasami też interakcja taka wspomagana jest dodatkowo wibracjami lub sygnalizacją dostępnymi diodami i wskaźnikami. Taka forma mieszana, udostępniająca wizualizację komunikatów głosowych, sama w sobie może również stanowić element edukacyjny – najłatwiej przecież nauczyć się języka na żywych przykładach i takie pozytywne efekty uboczne obserwuje się u dzieci grających w gry opracowane w języku angielskim, czy też bawiące się „rozmawiającymi” zabawkami.

Kanał zwrotny w projekcie interfejsu głosowego nie wymaga analizy semantycznej, może być zatem wdrażany jako stosunkowo proste, a zarazem atrakcyjne rozszerzenie istniejących systemów interaktywnych [4].

4. Zagrożenia skutecznej komunikacji za pomocą interfejsów głosowych

Do najważniejszych czynników utrudniających powszechne wdrożenie systemów opartych na komunikatach głosowych można zaliczyć ograniczoną jakość rozpoznania treści wypowiedzianych przez operatora, konieczność eliminacji zakłóceń, a także nieadekwatność tej formy komunikacji w niektórych kontekstach.

W sytuacjach, w których wymagana jest poufność informacji przekazywanej do systemu lub informacji zwrotnej, zastosowanie interfejsu mówionego będzie zawsze ryzykowne, nawet przy założeniu jakiejś formy fizycznej separacji operatora od otoczenia (np. przebywania w osobnym pomieszczeniu). Przykładami takich systemów wrażliwych mogą być transakcyjny system bankowości elektronicznej czy system rezerwacji usług medycznych. Niemniej jednak dla niektórych użytkowników krępujące, a nawet niedopuszczalne, może okazać się nawet upublicznianie listy zakupów. Wprowadzone w Polsce elektroniczne recepty wymagają podania pary PESEL – unikalny kod recepty, a ze względu na przepisy o ochronie danych osobowych i przekonania wielu osób w tym zakresie, trudno sobie wyobrazić system zmuszający użytkownika do przedyktowania numeru identyfikacyjnego PESEL głośno i wyraźnie w obecności wielu postronnych osób.

Algorytmy rozpoznawania mowy naturalnej mogą być implementowane w oparciu o uniwersalne wzorce dźwięków i głosek lub dostosowane do cech osobniczych konkretnego użytkownika w procesie treningu. To drugie podejście zapewnia lepszą jakość rozpoznania treści, lecz uniemożliwia wykorzystanie interfejsu głosowego z marszu, bez poświęcenia dodatkowego czasu i wysiłku. Nie jest więc ono w praktyce możliwe w przypadku rozwiązań dostępnych publicznie, jak przykładowo info-kioski czy terminale sprzedażowe. Z drugiej strony bez treningu systemu rozpoznawania identyfikacja operatora będzie często obciążona błędem uniemożliwiającym odróżnienie poszczególnych osób, a tym bardziej nie można budować w ten sposób rozwiązań wymagających ciągłej autoryzacji użytkownika.

Dodatkowym wyzwaniem jest implementacja systemów rozpoznawania mowy naturalnej w przypadku osób o niestandardowej charakterystyce głosu, a więc z wadami wymowy różnego rodzaju [5], posługujących się specyficznym akcentem czy dialektem, z głosem osłabionym różnorodnymi czynnikami – zaawansowanym wiekiem, urazami, zmianami zwyrodnieniowymi. Wypowiedzi takich osób zawierają często nienaturalne pauzy wynikające z trudności z oddechem czy artykułowaniem głosek. Podobne przeszkody mogą napotkać użytkownicy interfejsów głosowych w sytuacji chwilowej niedyspozycji, skutkującej zaburzeniami mowy, co może być związane z używkami, chorobami różnego rodzaju, jak np. cukrzyca, czy wreszcie z pospolitymi chorobami górnych dróg oddechowych.

Niezależnie od tego, czy obniżenie jakości rozpoznania głosu wynika ze stałej cechy użytkownika, czy jest tylko chwilowe, może niestety istotnie utrudnić skuteczne wdrożenie systemów z głosowym interfejsem użytkownika tam, gdzie są one najbardziej pożądane, m.in. w otoczeniu osób niepełnosprawnych i starszych użytkowników, dla których inne formy komunikacji człowiek-komputer (takie jak użycie ekranów dotykowych lub klawiatury z myszką) okazują się często zbyt obciążającym wyzwaniem.

W większości popularnych urządzeń rozpoznawanie poleceń głosowych wykorzystuje obecnie uniwersalne algorytmy i nie wymaga personalizacji w procesie treningu. W związku z tym, jeżeli sterowane urządzenia znajdują się w tym samym pomieszczeniu (lub ogólnie niedaleko siebie), może zachodzić sytuacja, w której nakładające się polecenia pochodzące od różnych operatorów będą powodować błędne działania. Takie przypadki miały miejsce w trakcie testów inteligentnych okularów Google Glass. Kontrowersje dotyczące takiego sprzętu są związane z tym, że osoby w otoczeniu ich użytkownika nie są w stanie zinterpretować, w jaki sposób używa on okularów, czy przerywają jego interakcję z urządzeniem ubieralnym, wtrącają się w nią, czy też ich wizerunek lub głos są właśnie w pewien sposób przetwarzane za pomocą kamery i mikrofonu wbudowanych w okulary. W ten sposób zagrożona jest zarówno prywatność użytkownika, jak i jego otoczenia [6].

W związku z tym, że interfejsy głosowe są projektowane nie tylko w kontekście zastosowań domowych i osobistych, ale również jako alternatywne rozwiązania w przemyśle, logistyce, transporcie należy brać pod uwagę różne rodzaje potencjalnych zakłóceń. Należą do nich szum tła otoczenia (który w środowisku przemysłowej hali produkcyjnej jest niemal nie do uniknięcia), a także inne dźwięki wydawane przez urządzenia, pojazdy i różne osoby postronne, które znajdują się często równie blisko, lub są równie głośne, jak operator. Zakłócenia takie mogą doprowadzić do przekłamania treści komunikatów głosowych, złej interpretacji poleceń głosowych czy też błędnego określenia źródła komunikacji, czyli użytkownika.

Informacja zwrotna generowana przez system oparty na komunikatach werbalnych również podlega potencjalnym zakłóceniom. Największym problemem wydaje się być ponownie nakładanie się głosu pochodzącego z wielu urządzeń w taki sposób, że użytkownik nie jest w stanie określić, czy komunikat dotyczy jego interakcji i czy jest generowany przez jego maszynę. Trafnym przykładem takiej sytuacji może być instalacja obok siebie wielu samoobsługowych terminali POS w sklepach. Z racji niewielkich odległości między terminalami informacja zwrotna może być nieczytelna dla odbiorcy. Nie zawiera ona przecież każdorazowo danych urządzenia, a jej znaczenie

zależy od kontekstu, np. zestawu rozliczanych towarów. Możliwy jest więc scenariusz, w którym nawet kilku użytkowników w podobnym stanie konwersacji jest wprowadzanych w błąd poprzez komunikat skierowany tylko do jednego z nich.

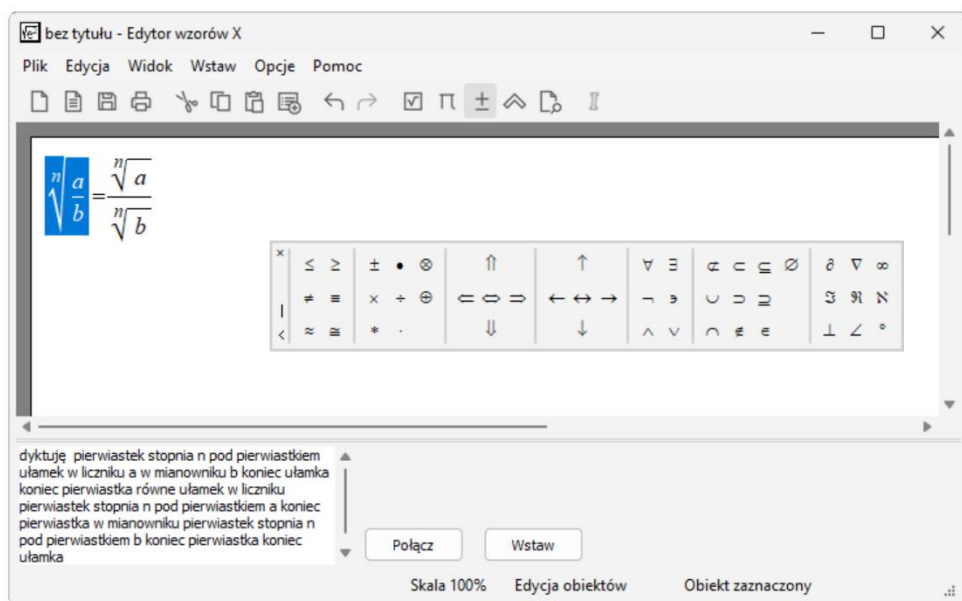
5. Projekt interfejsu głosowego – case study

Wdrożenie głosowej konwersacji między użytkownikiem a systemem komputerowym pozwala dostarczyć interakcję człowiek-komputer o unikalnej, odmiennej od klasycznej jakości [7]. Jest to szczególnie istotne dla systemów operujących na złożonych, skomplikowanych strukturach danych, takich jak przykładowo rozbudowane wyrażenia matematyczne. Standardowe podejście do edycji i przetwarzania takich danych wykorzystuje języki opisu dokumentu zbliżone do języków programowania (LaTeX, MathML), rozbudowane edytory wizualne wykorzystujące sterowanie za pomocą myszy komputerowej, a spotykane są również rozwiązania operujące na obrazie (np. wygenerowanym na ekranie dotykowym lub tablecie graficznym) i jego automatycznej analizie, a następnie konwersji do formy strukturalnej. Werbalizacja, czyli konwersja notacji strukturalnej na formę mówioną, a także możliwość dyktowania wyrażenia wraz z automatyczną interpretacją i konwersją do formy strukturalnej to dwa podstawowe elementy, pozwalające zaprojektować kompletny interfejs głosowy dla edycji wyrażeń matematycznych, kluczowy w rozwiązaniach e-learningowych dla osób niepełnosprawnych wzrokowo i ruchowo [8], a także jako narzędzie wspomagające wizualizację wzorów matematycznych w innych kontekstach.

Semantyka zapisu w notacji matematycznej określana jest nie tylko na podstawie znaczenia poszczególnych symboli i znaków specjalnych (np. sumy, całki), ale także ich wzajemnego ułożenia graficznego (np. granica po n vs. granica z n) oraz kontekstu lub obszaru tematycznego (np. wektor vs. punkt vs. macierz). Powyższe cechy powodują znaczącą trudność w automatyzacji odczytu i zapisu notacji matematycznej, opartej na zestawie komend głosowych [9]. W języku naturalnym ten sam wzór matematyczny może być przeczytany na wiele równorzędnych sposobów. Przykładowo ułamek $1/n$ możemy przeczytać jako: „1 podzielone na n ”, „w liczniku 1 w mianowniku n ”, „jedna n -ta”. Dlatego w procesie werbalizacji notacji matematycznej niezbędne jest ustalenie zestawu reguł i ich restrykcyjne przestrzeganie przy ustnym wprowadzaniu lub odczytywaniu wzorów.

Przykładem rozwiązania, w którym interfejs głosowy został zaimplementowany jako poszerzenie i uzupełnienie pośrednictwa użytkowego klasycznej aplikacji okienkowej dla systemu MS Windows, jest „Edytor wzorów” pozwalający na wizualną edycję, dyktowanie, wyszukiwanie i werbalizację zapisanych wyrażeń matematycznych [10].

W opisywanym narzędziu „Edytor wzorów” (rys. 1), zaprojektowano i zaimplementowano reguły werbalizacji dla podstawowych typów notacji matematycznej, takich jak: pierwiastki, ułamki, sumy, iloczyny, granice itp. Każda reguła składa się z komendy podstawowej (słowa kluczowego, np. „ułamek”) oraz komend wspomagających, które mają na celu odzwierciedlenie układu graficznego i współzależności między symbolami w ramach danej reguły (np. „w liczniku”, „w mianowniku”). W wielu przypadkach dla uzyskania jednoznaczności zapisu konieczne jest także wskazanie końca wyrażenia, którego dotyczy dana komenda (np. „koniec ułamek”).



Rysunek 1. Graficzne pośrednictwo użytkowe aplikacji „Edytor wzorów” wspierającej edycję wyrażeń matematycznych za pomocą interfejsu głosowego [opracowanie własne]

Rezultat dyktowania wzoru matematycznego przekazywany jest do edytora za pomocą mobilnej aplikacji pośredniczącej, której zadaniem jest odebranie strumienia dźwięku z urządzenia wejściowego i transkrypcja słów do formatu tekstowego. Główny proces przetwarzania wejściowego tekstu zaimplementowany jest w Edytorze w formie wspomnianych wyżej reguł. Mają one postać:

fraza_języka_naturalnego~format_wewnętrzny_EQED

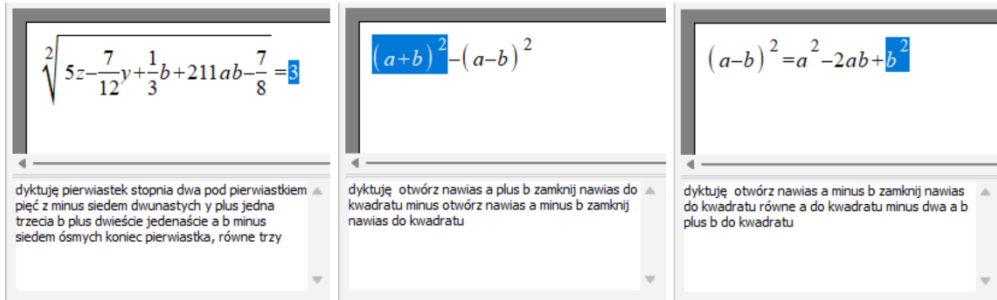
gdzie ~ (tylda) jest separatorem, lewa strona stanowi treść komendy, prawa strona – zapis komendy w formacie EQED – wewnętrznym formacie „Edytora wzorów”, będącym zbiorem makrodefinicji LaTeX. Przykładowe zapisy definicji reguł dla ułamków i pierwiastków:

```
ułamek~ \EQEDfrac
w liczniku~{
w mianowniku~}{
koniec ułamka~}
pierwiastek stopnia~\EQEDroot{
pierwiastek z~\EQEDroot{\EQEDplain{}}}{
pierwiastek~\EQEDroot{
stopnia z~}{
pod pierwiastkiem~}{
koniec pierwiastka~}
```

ilustrują proces translacji mowy naturalnej do notacji strukturalnej, opartej na zestawie komend-makrodefinicji zgodnych ze składnią LaTeX, z argumentami w nawiasach kłamrowych [11].

Symbole #, @ oraz \$ zostały dodatkowo wprowadzone dla obsługi podwyrażeń, które nie posiadają bezpośredniej komendy kończącej lub wymagają łączenia z innymi elementami w przypadku wyrażeń zagnieżdżonych (analogicznie jak w przypadku nawiasów).

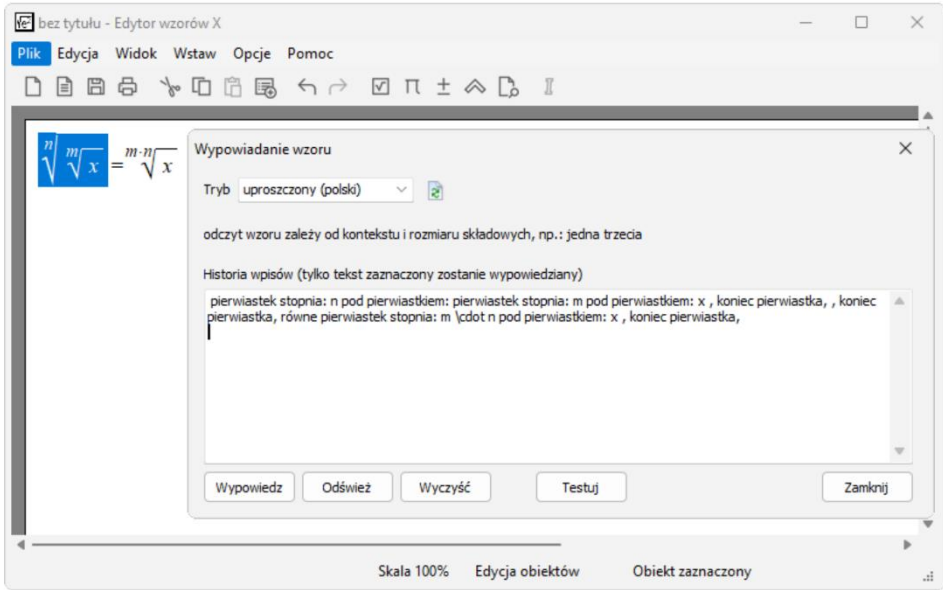
Opisane powyżej reguły pozwalają zapisać dyktowane wyrażenie matematyczne w wewnętrznym formacie edytora EQED, który następnie może być automatycznie konwertowany do wyjściowej postaci symbolicznej/wizualnej. Zaimplementowany translator notacji matematycznej sprawdza także poprawność użytych w strumieniu wejściowym komend i dokonuje korekty słów kluczowych poprzez skanowanie wstecz.



Rysunek 2. Przykłady wyrażeń matematycznych uzyskanych poprzez dyktowanie w aplikacji „Edytor wzorów” [opracowanie własne]

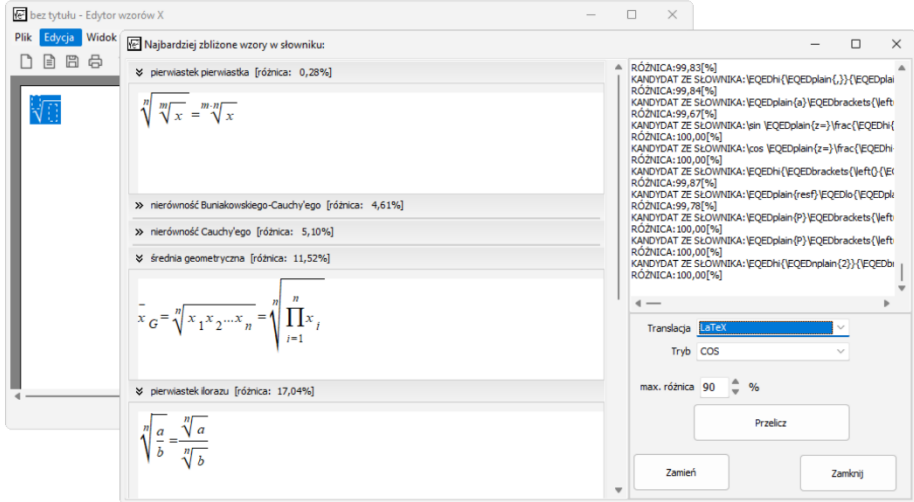
Tak przetworzony wzór może zostać wyeksportowany do jednego ze standardowych formatów zapisu notacji matematycznej (LaTeX, MathML) lub wyświetlony w zwizualizowanej postaci (rys. 2). Możliwa jest także konwersja odwrotna zapisanego w formacie EQED wyrażenia do postaci tekstowej, zwerbalizowanej i gotowej do odczytu (rys. 3).

Podobnie, możliwość przedyktowania wyrażenia matematycznego do systemu zapewniającego jego translację na mowę naturalną pozwala na przeprowadzenie wyszukiwania spośród zwerbalizowanych wzorów z repozytorium na podstawie cząstkowych informacji. Dzięki werbalizacji zwiększa się jakość wyszukiwania, które może być zaimplementowane z pomocą standardowych algorytmów wykorzystywanych do przybliżonego porównywania tekstów. Na rysunku 4 zaprezentowany jest moduł wyszukiwania oparty na takich założeniach [12] i rezultaty przykładowej kwerendy.



Rysunek 3. Werbalizacja wyrażenia matematycznego w programie „Edytor wzorów” [opracowanie własne]

W ten sposób „Edytor wzorów” może być wykorzystywany jako pomoc edukacyjna do automatycznego odczytywania cyfrowych źródeł zawierających notację matematyczną – podręczników, zbiorów zadań, treści wykładów itp.



Rysunek 4. „Edytor wzorów” – przeszukiwanie bazy danych wzorów na podstawie informacji cząstkowych [opracowanie własne]

6. Podsumowanie

Obserwowany w ostatnich latach dynamiczny rozwój głosowych interfejsów użytkownika oraz ich coraz to powszechniejsze stosowanie w urządzeniach codziennego użytku, otwiera nowe możliwości dla budowania narzędzi edukacyjnych i wspoma-

gających osoby niepełnosprawne ruchowo, wzrokowo bądź słuchowo. W rozdziale przedstawiono spektrum zastosowań interfejsów głosowych do budowania innowacyjnych rozwiązań interakcji człowiek-komputer, jak również najważniejsze zagrożenia dla efektywnej komunikacji za ich pomocą. W szczególności omówiono trudności i wyzwania w interpretacji werbalnej zapisu matematycznego i jego automatycznego przetwarzania. Na tym tle opisano założenia projektowe oraz implementację autorzkiego narzędzia służącego dyktowaniu i odsłuchiowaniu wyrażen matematycznych – programu „Edytor wzorów”.

Aplikacja umożliwi użytkownikowi zarówno wprowadzanie, jak i odtwarzanie zapisu matematycznego za pomocą intuicyjnych komend głosowych, dzięki czemu może ona stanowić uzupełnienie istniejących systemów e-learningowych o obsługę niewspieranych dotąd formuł zapisanych w notacji matematycznej. Istotnym obszarem wymagającym takich narzędzi jest edukacja osób z niepełnosprawnością, w szczególności w zakresie nauk matematycznych i inżynierskich, w których notacja symboliczna stanowi podstawę zapisu informacji. Również rozwój technologii e-learningowych i spopularyzowana w ostatnich latach zdalna forma edukacji wymuszają poszukiwanie rozwiązań ukierunkowanych na optymalizację metod komunikacji człowiek-maszyna w kierunku podobieństwa do interakcji międzyludzkich. Interfejsy głosowe wydają się być kluczowym elementem w tego typu komunikacji.

Literatura

1. Kowalski J., Jaskulska A., Skorupska K., Abramczuk K., Biele C., Kopeć W., Marasek K., *Older adults and voice interaction: A pilot study with Google Home*, [w:] *Extended Abstracts of the 2019 CHI Conference on human factors in computing systems*, s. 1-6.
2. Jaskulska A., Skorupska K., Karpowicz B., Biele C., Kowalski J., Kopeć W., *Exploration of Voice User Interfaces for Older Adults – A Pilot Study to Address Progressive Vision Loss*, [w:] *Conference on Multimedia, Interaction, Design and Innovation*, Springer, Cham, 2020, s. 159-168.
3. McTear M.F., Callejas Z., *Voice application development for Android*, Packt Publishing Ltd., 2013.
4. Pearl C., *Designing voice user interfaces: principles of conversational experiences*, O'Reilly Media, Inc., 2016.
5. Bętkowska Cavalcante A.B., Lorens L., *Use case: a mobile speech assistant for people with speech disorders*, [w:] *Proceedings of the 7th Language & Technology Conference*, 2015, s. 192-197.
6. Due B., *Challenges with Google Glass in social interaction*, [w:] *Proceedings of the 4th Participatory Innovation Conference PIN-C*, 2015, s. 440-448.
7. Lee H., *Voice user interface projects: Build voice-enabled applications using Dialogflow for Google home and Alexa skills kit for Amazon Echo*, Packt Publishing Ltd., 2018.
8. Attanayake D., Denholm-Price J., Hunter G., Pflugel E., Wigmore A., *Speech interfaces for mathematics: Opportunities and limitations for visually impaired learners*, [w:] *IMA International Conference on Barriers and Enablers to Learning Maths: Enhancing Learning and Teaching for All Learners*, 2015, s. 1-8.
9. Cuartero-Olivera J., Hunter G., Perez-Navarro A., *Reading and writing mathematical notation in e-learning environments*, eLearn Center Research Paper Series, no. 4, 2012, s. 11-20.
10. Bier A., Sroczyński Z., *Rule based intelligent system verbalizing mathematical notation*, *Multimedia Tools and Applications*, 78(19), 2019, s. 28089-28110.

11. Bier A., Sroczyński Z., *Rule-based intelligent system for dictating mathematical notation in Polish*, [w:] Mauri J.M. (red.), *ACHI 2020: The Thirteenth International Conference on Advances in Computer-Human Interactions*, Valencia, Spain, 21-25 November 2020, ISBN 978-1-61208-761-0, s. 352-356.
12. Bier A., Sroczyński Z., *Towards semantic search for mathematical notation*, [w:] *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, IEEE, 2018, s. 465-469.

Głosowe interfejsy użytkownika: zastosowania, projektowanie, zagrożenia

Streszczenie

Dynamiczny rozwój wydajności sprzętu komputerowego spowodował, że w ostatnich latach rozpoznawanie i generowanie mowy naturalnej stało się dostępne zarówno dla komputerów biurowych (laptopów czy stacjonarnych), jak i popularnych urządzeń mobilnych. Syntetyzowanie mowy naturalnej osiągnęło już jakość, dzięki której rezultat jest trudny do odróżnienia od nagrania ludzkiego lektora. Natomiast rozpoznawanie mowy może być prowadzone z dokładnością pozwalającą na efektywne dyktowanie tekstu. W związku z powyższym w coraz szerszym zakresie obserwujemy próby wdrażania interfejsów użytkownika opartych na głosie zamiast standardowych urządzeń wejścia/wyjścia. W większości przypadków opierają się one na odpowiednim słowniku poleceń. W rozdziale jako przykład takiego interfejsu przedstawiono środowisko edukacyjne wspierające dyktowanie i edycję złożonych dokumentów matematycznych, stanowiące efektywne narzędzie komunikacji człowiek-komputer z wykorzystaniem mowy naturalnej.

Słowa kluczowe: głosowe interfejsy użytkownika, interakcja człowiek-komputer, doświadczenie użytkownika, werbalizacja

Voice user interfaces: applications, design, threats

Abstract

The dynamic increase of computer hardware efficiency has made natural speech recognition and generation available for both office computers (laptops or desktops) and popular mobile devices. Synthesizing natural speech has already reached a quality that makes the result difficult to distinguish from a human voice recording. On the other hand, speech recognition can be carried out with an accuracy that allows for effective dictation of the text. Therefore, we observe multiple attempts to implement voice-based user interfaces instead of standard input / output devices. In most cases, they are based on the appropriate command dictionary. As an example of such an interface, the chapter presents an educational environment supporting the dictation and editing of complex mathematical documents, which is an effective tool for human-computer communication using natural speech.

Keywords: voice user interfaces, human-computer interaction, user experience, verbalization

Zastosowania i nowe rozwiązania dla głębokich sieci neuronowych

1. Wprowadzenie

Uczenie głębokie bazuje na warstwach sieci neuronowych, które są algorytmami modelowanymi w przybliżeniu do sposobu działania ludzkiego mózgu. Trening z dużą ilością danych konfiguruje neurony w sieci neuronowej, w wyniku czego powstaje model uczenia, który po przeszkoleniu przetwarza nowe dane. Modele uczenia głębokiego pobierają informacje z wielu źródeł danych i analizują te dane w czasie rzeczywistym, bez konieczności interwencji ze strony człowieka. W procesie uczenia tego typu procesory graficzne (GPU, ang. *Graphics Processing Unit*) są zoptymalizowane pod kątem modeli szkoleń, ponieważ mogą przetwarzać wiele obliczeń jednocześnie. Uczenie to napędza wiele technologii sztucznej inteligencji (AI, ang. *Artificial Intelligence*), które mogą usprawnić automatyzację i zadania analityczne.

W ciągu ostatniej dekady głębokie sieci neuronowe (DBN, ang. *Deep Belief Network*) stały się potężnym narzędziem w dziedzinie uczenia maszynowego i są z powodzeniem stosowane do rozwiązywania wielu problemów sztucznej inteligencji, takich jak rozpoznawanie obrazów, rozpoznawanie mowy, widzenie komputerowe, przetwarzania języka naturalnego, czy wizualizacja danych [1-10]. Głębokie sieci neuronowe są tematem rozważań wielu prac badawczych [1-10]. Składają się z ukrytych warstw [1] i mogą przeprowadzać głęboką hierarchiczną transformację danych wejściowych, w wyniku czego stwierdzono, że mają lepszą wydajność i większą moc reprezentacyjną niż tradycyjne sieci neuronowe, znane również jako sztuczne sieci neuronowe (ANN, ang. *Artificial Neural Network*) lub symulowane sieci neuronowe (SNN, ang. *Spiking Neural Network*).

Istotnym problemem jest uczenie głębokiej sieci neuronowej, ponieważ uczenie takiej sieci jest znacznie bardziej skomplikowane w porównaniu do tradycyjnych sieci neuronowych. Wynika to z problemu znikającego gradientu, słabych minimów lokalnych i problemu niestabilnego gradientu. Dlatego opracowano wiele technik głębokiego uczenia się, które pozwalają przezwyciężyć ograniczenia konwencjonalnych podejść szkoleniowych [1, 2]. Od 2006 roku zaproponowano nienadzorowane wstępne szkolenie głębokich sieci neuronowych [3]. W tym przypadku uczenie sieci neuronowych składa się z dwóch etapów: wstępnego uczenia sieci neuronowej przy użyciu podejścia warstwowego oraz dostrajania wszystkich parametrów sieci neuronowej przy użyciu

¹ v.golovko@dyd.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl.

² m.chodyka@dyd.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl.

³ p.lichograj@dyd.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl.

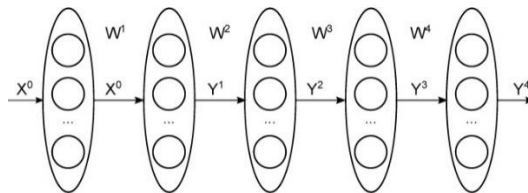
⁴ grzegorski22062@stud.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl.

algorytmu propagacji wstecznej lub algorytmu wake-sleep [4-7]. Trening wstępny sieci neuronowej opiera się na maszynie Boltzmanna (RBM, ang. *Restricted Boltzman Machines*) albo na podejściu autokoderowym [7]. Obecnie do uczenia głębokich sieci neuronowych w sposób nadzorowany wykorzystywana jest funkcja aktywacji stochastycznego gradientu (SGD, ang. *Stochastic Gradient Descent*) z rektyfikowaną jednostką liniową (ReLU, ang. *Rectified Linear Unit*) [1].

Istnieją różne typy głębokich sieci neuronowych, a mianowicie głębokie sieci przekonań, głębokie konwolucyjne sieci neuronowe, głębokie rekurencyjne sieci neuronowe, głębokie autokodery i inne. Ten artykuł zawiera przegląd sieci neuronowych tego rodzaju i głębokiego uczenia się. Omówiono różne techniki głębokiego uczenia się i nowe trendy obecnie obowiązujące. Zaprezentowano nową technikę do uczenia sieci neuronowych, opartą na ograniczonej maszynie Boltzmanna. Podejście to, w przeciwieństwie do modelu bazującego na energii, opiera się na minimalizacji średniokwadratowego błędu rekonstrukcji w ukrytych i rekonstruowanych warstwach RBM. Pokazano, że klasyczne równania do treningu RBM są szczególnym przypadkiem proponowanej techniki. Eksperymenty pokazały duży potencjał głębokich sieci neuronowych w rzeczywistych zastosowaniach. Pozostała część artykułu jest zorganizowana w następujący sposób. Rozdział 2 przedstawia współczesne techniki głębokiego uczenia się. Rozdział 3 opisuje konwencjonalne podejście do ograniczonego szkolenia maszynowego Boltzmanna. Rozdział 4 dotyczy funkcji aktywacji stochastycznego gradientu. W rozdziale 5 zaproponowano autorskie podejście do wnioskowania o regułach szkolenia RBM. Rozdziały 6 i 7 przedstawiają obecny paradygmat uczenia się, wnioski i przyszłość głębokiego uczenia się.

2. Współczesne techniki głębokiego uczenia się

Konwencjonalne podejście do uczenia maszynowego RBM [3-8] opiera się na modelu opartym na energii. Jak już wspomniano, DBN składa się z wielu warstw i może wykonywać głęboką hierarchiczną reprezentację danych wejściowych, jak pokazano na rysunku 1.



Rysunek 1. Przykład głębokiej sieci neuronowej [opracowanie własne]

J – ta jednostka wyjściowa dla k – tej warstwy jest dana przez:

$$y_j^k = F(S_j^k), \quad (1)$$

$$S_j^k = \sum_{i=1} w_{ij}^k y_i^{k-1} + T_j^k \quad (2)$$

gdzie F jest funkcją aktywacji, S_j^k jest sumą ważoną j – tej jednostki, w_{ij}^k jest wagą od i – tej jednostki ($k-1$) – tej warstwy do j – tej jednostki k – tej warstwy, T_j^k jest progmem j – tej jednostki.

Na pierwszą warstwę

$$y_i^0 = x_i \quad (3)$$

W powszechnym przypadku można to zapisać:

$$Y^k = F(S^k) = F(W^k Y^{k-1} + T^k) \quad (4)$$

gdzie W jest macierzą wag, Y^{k-1} jest wektorem wyjściowym dla $(k-1)$ – tej warstwy, T^k jest wektorem progowym.

Należy również zauważyć, że wyjście DBN jest często definiowane za pomocą funkcji *softmax*:

$$y_j^F = \text{soft max}(S_j) = \frac{e^{S_j}}{\sum_l e^{S_l}} \quad (5)$$

Istnieją następujące techniki uczenia głębokich sieci neuronowych: uczenie z przed-treningiem oraz podejście stochastycznego gradientu SGD z funkcją przenoszenia rektyfikowanej jednostki liniowej ReLU [1].

Nauka z przygotowaniem wstępnym składa się z dwóch etapów. Pierwszym etapem jest wstępne uczenie sieci neuronowej przy użyciu zachłannego podejścia warstwowego. Procedura ta rozpoczyna się od pierwszej warstwy i jest wykonywana bez nadzoru. Drugi to dostrajanie wszystkich parametrów sieci neuronowej za pomocą algorytmu wstecznej propagacji lub uśpienia [2-10].

Trening z podejściem stochastycznego gradientu to uczenie online lub *mini-batch* przy użyciu konwencjonalnego algorytmu wstecznej propagacji. Zastosowanie funkcji przenoszenia rektyfikowanej jednostki liniowej ReLU może pomóc uniknąć problemu znikającego gradientu, złych minimów lokalnych i problemu niestabilnego gradientu ze względu na większą liniowość tego rodzaju funkcji aktywacji.

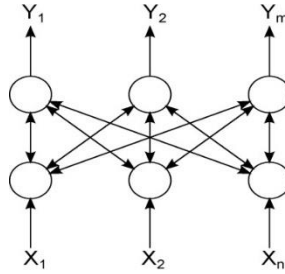
Tak więc obecnie istnieją dwie główne techniki uczenia głębokich sieci neuronowych, a mianowicie uczenie ze wstępnym szkoleniem oraz metoda stochastycznego spadku gradientu z rektyfikowaną funkcją aktywacji jednostek liniowych. Wybór podejścia zależy od wymiaru zbioru danych uczących.

3. Ograniczona maszyna Boltzmana

Ważnym etapem uczenia głębokich sieci neuronowych jest wstępne szkolenie każdej warstwy DBN. Trening wstępny DBN jest oparty na ograniczonej maszynie Boltzmana RBM lub na podejściu z automatycznym koderem [9]. Zgodnie z zachłanną procedurą uczenia warstwowego, na początku pierwsza warstwa DBN jest trenowana przy użyciu reguł uczenia RBM lub autokodera. Następnie szkolone są kolejne warstwy jedna po drugiej. W rezultacie uzyskuje się dobrą inicjalizację sieci neuronowej, a następnie można zastosować algorytm wstecznej propagacji lub algorytm *wake-sleep* do dostrajania parametrów całej sieci neuronowej.

W tej sekcji przedstawiona została technika przedtreningowa DBN oparta na ograniczonej maszynie Boltzmana. Głębokie sieci przekonań można przedstawić jako zbiór ograniczonych maszyn Boltzmana. Dlatego w tym przypadku RBM jest głównym budulcem głębokich sieci przekonań. Tradycyjne podejście do treningu RBM

opiera się na modelu energetycznym i regułach treningowych [2], które uwzględniają jedynie liniowy charakter jednostek neuronalnych. Ograniczona maszyna Boltzmanna jest w rzeczywistości szczególnym przypadkiem maszyny Boltzmanna, w której neurony tej samej warstwy są od siebie niezależne. Ograniczona maszyna Boltzmanna składa się z dwóch warstw jednostek: widocznej i ukrytej (rys. 2). Może ona reprezentować dowolny rozkład dyskretny, jeśli użyto wystarczającej liczby jednostek ukrytych [6]. Ukryte jednostki RBM to detektory cech, które wychwytyują prawidłowości danych wejściowych. Stany wszystkich jednostek uzyskuje się poprzez rozkład prawdopodobieństwa.



Rysunek 2. Ograniczona maszyna Boltzmanna [opracowanie własne]

Warstwy jednostek neuronowych są połączone dwukierunkowymi wagami W . Jednostki binarne są bardzo często używane [2]. RBM jest stochastyczną siecią neuronową, a stany jednostek widocznych i ukrytych są definiowane za pomocą probabilistycznej wersji funkcji aktywacji sigmoidy:

$$p(y_j | x) = \frac{1}{1 + e^{-S_j}}, S_j = \sum_i^n w_{ij} x_i + T_j \quad (6)$$

$$p(x_i | y) = \frac{1}{1 + e^{-S_i}}, S_i = \sum_j^m w_{ij} y_j + T_i \quad (7)$$

Kluczową ideą treningu RBM jest jak najdokładniejsze odtworzenie rozkładu danych wejściowych przy użyciu stanów jednostek ukrytych. Jest to równoznaczne z maksymalizacją prawdopodobieństwa rozkładu danych $P(x)$ poprzez modyfikację wag synaptycznych za pomocą gradientu logarytmu prawdopodobieństwa danych wejściowych [2]. Następnie modyfikacja wag synaptycznych jest określona przez:

$$\omega_{ij}(t + 1) = \omega_{ij}(t) + \alpha \frac{\partial \ln P(x)}{\partial \omega_{ij}(t)}$$

Korzystając z tego podejścia, Hinton zaproponował zastosowanie techniki kontrastywnej dywergencji (CD) do uczenia się RBM [2]. Opiera się ona na samplowaniu Gibbsa. W tym przypadku pierwszy człon reguły uczącej oznacza rozkład danych w czasie $t = 0$, a drugi człon to modelowy rozkład zrekonstruowanych stanów w kroku $t = k$. W rezultacie możemy uzyskać zasady szkolenia RBM. W przypadku CD-1 reguła ucząca jest zdefiniowana jako:

$$\begin{aligned}
 w_{ij}(t+1) &= w_{ij}(t) + \alpha(x_i(0)y_j(0) - x_i(1)y_j(1)) \\
 T_i(t+1) &= T_i(t) + \alpha(x_i(0) - x_i(1)) \\
 T_j(t+1) &= T_j(t) + \alpha(y_j(0) - y_j(1)).
 \end{aligned}
 \tag{8}$$

Jeśli zostanie użyte CD-k wówczas:

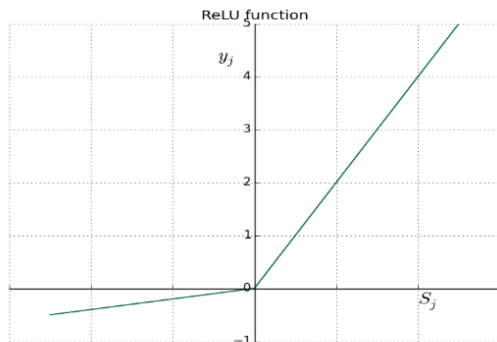
$$\begin{aligned}
 w_{ij}(t+1) &= w_{ij}(t) + \alpha(x_i(0)y_j(0) - x_i(k)y_j(k)) \\
 T_i(t+1) &= T_i(t) + \alpha(x_i(0) - x_i(k)) \\
 T_j(t+1) &= T_j(t) + \alpha(y_j(0) - y_j(k)).
 \end{aligned}
 \tag{9}$$

Jak widać z powyższych równań, zasady uczenia dla RBM zasadniczo minimalizują różnicę między oryginalnymi danymi a zsyntetyzowanymi próbkami z modelu. Zsyntetyzowane dane można uzyskać za pomocą algorytmu próbkowania Gibbsa (ang. *Gibbs sampler, Glauber dynamics*), który jest używany do symulacji rozkładu prawdopodobieństwa na przestrzeni stanów.

Uczenie RBM polega na przedstawieniu próbki uczącej jednostkom widocznym, a następnie zastosowaniu procedury CD-n do obliczenia stanów binarnych jednostek ukrytych, próbkowaniu jednostek widocznych (stany zrekonstruowane). Po wykonaniu iteracji wagi i odchylenia ograniczonej maszyny Boltzmanna są aktualizowane. Następnie układany jest na kolejnej ukrytej warstwie, aby wytrenować nowy RBM. Podejście to stosuje się do wszystkich warstw głębokich sieci przekonań (chciwy trening warstwowy). W wyniku tego wstępnego uczenia bez nadzoru można uzyskać dobrą inicjalizację sieci neuronowej. Na koniec przeprowadzane jest nadzorowane dostrajanie całej sieci neuronowej.

4. Funkcja aktywacji stochastycznego gradientu z rektyfikowaną jednostką liniową

Jak wspomniano wcześniej, uczenie się przy użyciu stochastycznego podejścia gradientowego reprezentuje uczenie online lub *mini-batch* przy użyciu konwencjonalnego algorytmu propagacji wstecznej. W tym przypadku wykorzystywana jest funkcja aktywacji (rys. 3).



Rysunek 3. Funkcja aktywacji ReLU [opracowanie własne]

Zgodnie z tą funkcją wyjście neuronu jest określone wzorem:

$$y_j = F(S_j) = \begin{cases} S_j, S_j > 0 \\ kS_j, S_j \leq 0 \end{cases}$$

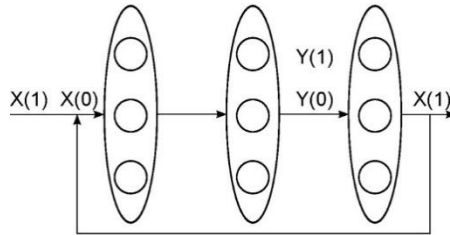
gdzie $k = 0$ lub inna mała wartość, na przykład $k = 0,01$ lub $0,001$.

Zastosowanie funkcji przenoszenia rektyfikowanej jednostki liniowej ReLU może przewyżżyć ograniczenia standardowego podejścia wstecznej propagacji [9, 11], a mianowicie problem znikającego gradientu, słabe minima lokalne i problem niestabilnego gradientu.

5. Autorskie rozwiązanie

W tym rozdziale przedstawione zostało autorskie podejście związane ze szkoleniem RBM, które cały czas jest rozwijane. Rozwiązanie zostało oparte na minimalizacji średniokwadratowego błędu rekonstrukcji, uzyskanego za pomocą prostych iteracji próbkowania Gibbsa. W porównaniu z tradycyjną metodą energetyczną, która opiera się na liniowej reprezentacji jednostek neuronowych, zaproponowane podejście pozwoliło uwzględnić nieliniowy charakter jednostek neuronowych [8-10].

Opierając się na ograniczonej maszynie Boltzmann, reprezentacja RBM jest możliwa przy użyciu trzech warstw (widocznej, ukrytej i widocznej) [8, 9], jak pokazano na rysunku 4.



Rysunek 4. Zmodyfikowana ograniczona maszyna Boltzmana [opracowanie własne]

Próbkowanie Gibbsa będzie się składać z następującej procedury. Niech $x(0)$ będzie danymi wejściowymi, które przechodzą do warstwy widocznej w czasie 0. Wtedy wyjście warstwy ukrytej definiuje się następująco:

$$y_j(0) = F(S_j(0)) \quad (10)$$

$$S_j(0) = \sum_i w_{ij} x_i(0) + T_j \quad (11)$$

Warstwa odwrotna rekonstruuje dane z warstwy ukrytej. W rezultacie można otrzymać $x(1)$ w czasie 1:

$$x_i(1) = F(S_i(1)) \quad (12)$$

$$S_i(1) = \sum_j w_{ij} y_j(0) + T_i \quad (13)$$

Następnie x (1) wchodzi do warstwy widocznej i można uzyskać wynik warstwy ukrytej w następujący sposób:

$$y_j(1) = F(S_j(1)) \quad (14)$$

$$S_j(1) = \sum_i w_{ij} x_i(1) + T_j \quad (15)$$

Kontynuując dany proces, można otrzymać na kroku k :

$$y_j(k) = F(S_j(k)), S_j(k) = \sum_i \omega_{ij} x_i(k) + T_j$$

$$x_i(k) = F(S_i(k)), S_i(k) = \sum_j w_{ij} y_j(k-1) + T_i$$

Celem uczenia tej sieci neuronowej jest zminimalizowanie średniokwadratowego błędu rekonstrukcji (ang. *Mean Squared Error*, MSE) w warstwie ukrytej i widocznej. W przypadku CD-k średniokwadratowy błąd rekonstrukcji definiuje się jako:

$$\begin{aligned} E_s &= \frac{1}{2} \sum_{l=1}^L \sum_{j=1}^m \sum_{p=1}^k (y_j^l(p) - y_j^l(p-1))^2 + \\ &+ \frac{1}{2} \sum_{l=1}^L \sum_{i=1}^n \sum_{p=1}^k (x_i^l(p) - x_i^l(p-1))^2 \end{aligned} \quad (16)$$

W przypadku CD-1 możemy pisać:

$$\begin{aligned} E_s &= \frac{1}{2} \sum_{l=1}^L \sum_{j=1}^m (y_j^l(1) - y_j^l(0))^2 + \\ &+ \frac{1}{2} \sum_{l=1}^L \sum_{i=1}^n (x_i^l(1) - x_i^l(0))^2 \end{aligned} \quad (17)$$

gdzie L to liczba wzorców treningowych.

Twierdzenie 1. Maksymalizacja logarytmicznego rozkładu wiarygodności danych wejściowych $P(x)$ w przestrzeni maszyny Boltzmann z ograniczonymi wagami synaptycznymi jest równoważna minimalizacji średniokwadratowego błędu rekonstrukcji w tej samej przestrzeni przy użyciu neuronów liniowych w RBM.

Teoria ta mówi, że jeśli użyje się funkcji aktywacji tożsamości dla jednostek RBM, to reguła uczenia CD-k dla RBM w celu zminimalizowania średniokwadratowego błędu rekonstrukcji (16) będzie następująca:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(x_i(0)y_j(0) - x_i(k)y_j(k)),$$

$$T_j(t+1) = T_j(t) + \alpha(y_j(0) - y_j(k)),$$

$$T_i(t+1) = T_i(t) + \alpha(x_i(0) - x_i(k))$$

Jak widać, ostatnie równania są identyczne z konwencjonalnymi regułami treningu RBM. Zatem konwencjonalne zasady szkolenia RBM są liniowe. Dlatego taką maszynę można nazywać liniowymi RBM.

Wniosek 1. Liniowa ograniczona maszyna Boltzmanna ze szkoleniowego punktu widzenia jest równoważna liniowej sieci neuronowej PCA (autoasocjacyjnej), jeśli podczas uczenia zastosowane jest próbkowanie Gibbsa.

Wniosek 2. Reguła ucząca dla nieliniowej ograniczonej maszyny Boltzmanna w przypadku CD-k jest zdefiniowana jako:

$$\begin{aligned}
 w_{ij}(t+1) &= w_{ij}(t) - \\
 &\alpha \left(\sum_{p=1}^k (y_j(p) - y_j(p-1)) x_i(p) F'(S_j(p)) + \right. \\
 &\left. (x_i(p) - x_i(p-1)) y_j(p-1) F'(S_i(p)) \right) \\
 T_j(t+1) &= T_j(t) - \\
 &- \alpha \left(\sum_{p=1}^k (y_j(p) - y_j(p-1)) F'(S_j(p)) \right), \\
 T_i(t+1) &= T_i(t) - \\
 &- \alpha \left(\sum_{p=1}^k (x_i(p) - x_i(p-1)) F'(S_i(p)) \right)
 \end{aligned}$$

Jak więc widać, klasyczne równania treningu RBM są szczególnym przypadkiem proponowanej techniki.

Twierdzenie 2. Maksymalizacja logarytmicznego rozkładu danych wejściowych $P(x)$ w przestrzeni ograniczonej maszyny Boltzmanna wag synaptycznych jest równoznaczna z minimalizacją dywergencji KL (entropii krzyżowej).

Wniosek 3. Maksymalizacja logarytmicznego rozkładu wiarygodności danych wejściowych $P(x)$ w przestrzeni maszyny Boltzmanna z ograniczonymi wagami synaptycznymi jest równoważna minimalizacji entropii krzyżowej i minimalizacji średniokwadratowego błędu rekonstrukcji przy użyciu neuronów liniowych w RBM.

$$\max(\ln P(x)) = \min(KL) = \min(E_s)$$

Zatem tradycyjna metoda bazująca na energii opiera się na maksymalizacji logarytmicznego rozkładu danych wejściowych i prowadzi do liniowej reprezentacji jednostek neuronowych z punktu widzenia uczenia. Proponowane podejście opiera się na minimalizacji średniokwadratowego błędu rekonstrukcji, który możemy uzyskać za pomocą prostych iteracji próbkowania Gibbsa i prowadzi do nieliniowej reprezentacji neuronów.

Jeśli zestaw danych treningowych jest duży, wówczas do głębokiego uczenia sieci neuronowych stosuje się SGD z ReLU. W przeciwnym razie stosuje się wstępne szkolenie i dostrajanie. Jak stwierdzono w [1]: W przypadku mniejszych zestawów danych wstępne szkolenie bez nadzoru pomaga zapobiegać nadmiernemu dopasowaniu,

prowadząc do znacznie lepszego uogólnienia, gdy liczba oznaczonych przykładów jest niewielka lub w ustawieniach przesyłania, w których jest wiele przykładów dla niektórych zadania źródłowe, ale bardzo niewiele w przypadku niektórych zadań „docelowych”. Po przywróceniu głębokiego uczenia się okazało się, że etap przedtreningowy był potrzebny tylko w przypadku małych zestawów danych.

6. Wnioski

Tradycyjna metoda energetyczna opiera się na maksymalizacji logarytmicznego rozkładu danych wejściowych i prowadzi do liniowej reprezentacji jednostek neuronowych z punktu widzenia uczenia. Przedstawiono, że maksymalizacja logarytmicznego rozkładu wiarygodności danych wejściowych $P(x)$ w przestrzeni ograniczonej wagami synaptycznymi maszyny Boltzmanna jest równoważna minimalizacji entropii krzyżowej i minimalizacji średniokwadratowego błędu rekonstrukcji przy użyciu neuronów liniowych w RBM. Proponowane podejście opiera się na minimalizacji średniokwadratowego błędu rekonstrukcji, który można uzyskać za pomocą prostych iteracji próbkowania Gibbsa i prowadzi do nieliniowej reprezentacji neuronów. Konwencjonalne równania do treningu RBM są szczególnym przypadkiem proponowanej techniki, gdy używana jest funkcja przenoszenia tożsamości. Przedstawione wnioski dotyczą stwierdzenia, iż liniowo ograniczona maszyna Boltzmanna ze szkoleniowego punktu widzenia jest równoważna liniowej sieci neuronowej PCA (autoasocjacyjnej), jeśli podczas uczenia zastosowane jest próbkowanie Gibbsa. Dodatkowo przedstawiono, iż tradycyjna metoda oparta na energii opiera się na maksymalizacji logarytmicznego rozkładu danych wejściowych i prowadzi do liniowej reprezentacji jednostek neuronowych z punktu widzenia uczenia.

Literatura

1. LeCun Y., Bengio Y., Hinton G.E., *Deep Learning Nature*, 521, 2015, s. 436-444.
2. Golovko V.A., *New Technique for Restricted Boltzmann Machine Learning*, Golovko V.A., Kroshchanka A., Turchenko V., Jankowski S., Treadwell D., *Proceedings of the 8th IEEE International Conference IDAACS*, 2015, s. 182-186.
3. Hinton G., *Training products of experts by minimizing contrastive divergence*, Neural Computation, 14, 2002, s. 1771-1800.
4. Hinton G., Salakhutdinov R., *Reducing the dimensionality of data with neural networks*, Science, 313(5786), 2006, s. 504-507.
5. Hinton G.E., *A practical guide to training restricted Boltzmann machines*, Tech. Rep. 2010-000, Machine Learning Group, University of Toronto, Toronto 2010.
6. Hinton G.E., Osindero S., Teh Y., *A fast learning algorithm for deep belief nets*, Neural Computation, 18, 2006, s. 1527-1554.
7. Bengio Y., *Learning deep architectures for AI. Foundations and Trends in Machine Learning*, 2(1), 2009, s. 1-127.
8. Bengio Y., Lamblin P., Popovici D., Larochelle H., *Greedy layer-wise training of deep networks*, [w:] Scholkopf B., Platt J.C., Hoffman T. (red.), *Advances in neural information processing systems*, MA: MIT Press, 11, Cambridge 2007, s. 153-160.
9. Golovko V.A., *From multilayer perceptron to deep belief neural networks: training paradigms and application*, Lectures on neuroinformatics, Moscow 2015, s. 47-84.
10. Golovko V.A., *Learning Technique for Deep Belief Neural Networks*, Golovko V.A., Kroshchanka A., Rubanau U., Jankowski S., in book *Neural Networks and Artificial Intelligence – Springer, Communication in Computer and Information Science*, 440, 2014, s. 136-146.

Zastosowania i nowe rozwiązania dla głębokich sieci neuronowych

Streszczenie

W ciągu ostatniej dekady głębokie sieci neuronowe stały się rewolucyjną techniką w dziedzinie sztucznej inteligencji i uczenia maszynowego. W ogólnym przypadku głęboka sieć neuronowa składa się z wielu warstw jednostek neuronowych i może realizować głęboką hierarchiczną reprezentację swoich danych wejściowych. Pierwsza warstwa wyodrębnia funkcje niskiego poziomu; druga warstwa wykrywa cechy wyższego poziomu, w wyniku czego głęboka sieć neuronowa dokonuje głębokiej nieliniowej transformacji danych wejściowych na bardziej abstrakcyjny poziom reprezentacji. Ten artykuł zawiera przegląd głębokich sieci neuronowych i głębokiego uczenia. Omówiono różne techniki głębokiego uczenia się obecnie znane i nowe trendy w tym zakresie.

Słowa kluczowe: maszyna Boltzmanna, głębokie sieci neuronowe, głębokie uczenie

Applications and new solutions for deep neural networks

Abstract

Over the last decade the deep neural networks are the revolutionary technique in the domain of artificial intelligence and machine learning. In the general case a deep neural network consists of multiple layers of neural units and can accomplish a deep hierarchical representation of their input data. The first layer extracts low-level features; the second layer detects higher level features, and as a result the deep neural network performs deep non-linear transformation of input data into more abstract level of representation. This paper provides an overview of deep neural networks and deep learning. Different deep learning techniques, including well-known and new approaches are discussed.

Keywords: Restricted Boltzmann machine, deep neural networks, deep learning

Metoda wykrywania napadów padaczkowych za pomocą fal EEG

1. Wprowadzenie

Neuronauka to nauka o procesach chemicznych, biologicznych i anatomicznych, które mają znaczący wpływ na zachowanie i funkcje mózgu. Łączy różne interdyscyplinarne dziedziny, takie jak medycyna, chemia, psychologia, biologia molekularna, anatomia, czy fizyka. Pozwala na badanie układu nerwowego i wpływu nerwów na zachowanie przy użyciu szerokiego zakresu podejść naukowych. Zwana jest także nauką o neuronach, gdyż stara się zrozumieć, w jaki sposób układ nerwowy funkcjonuje u osób zdrowych, jak i u osób z zaburzeniami mózgu, zaburzeniami psychicznymi lub neurorozwojowymi. Skupia się przede wszystkim na budowie i rozwoju ośrodkowego układu nerwowego. Z tego powodu badania neurologiczne często koncentrują się na tym, jak mózg wpływa na zachowanie i funkcje poznawcze. Wraz z rosnącą liczbą urządzeń, takich jak elektroencefalogram (EEG, ang. *Electroencephalography*), rezonans magnetyczny (MRI, ang. *Magnetic Resonance Imaging*) czy technologii obrazowania 3D, dziedzina ta pomaga rozszyfrować złożone działanie mózgu. Przykładem tego typu działań jest wykrywanie napadów padaczkowych i wykorzystanie do ich diagnostyki i terapii badania elektroencefalograficznego.

Napady padaczkowe będące powszechnym schorzeniem neurologicznym dotyczą około 1% światowej populacji [1]. Charakterystyka metod ich wykrywania za pomocą fal EEG jest szeroko opisana w literaturze, a EEG skóry głowy jest najczęściej stosowanym narzędziem diagnostycznym w padaczce. Wykrywanie napadów padaczkowych, jak również wykrywanie aktywności padaczkowopodobnej czy międzynapadowej, jest istotną częścią codziennego postępowania z pacjentami z padaczką. Szereg metod dotyczących detekcji i predykcji napadów stosowane jest od lat 70. ubiegłego wieku. W latach 80.-90. ubiegłego wieku stosowano do predykcji napadów parametrów opartych o metody nieliniowe analizy sygnałów, takich jak wymiar korelacyjny i wykładniki Lyapunowa [2]. Stosowanie tych metod było wówczas modne, jednak okazało się, że nie do końca spełnia oczekiwania badaczy. Na przełomie wieków pojawiły się prace prezentujące krytyczny pogląd na temat ich stosowania np. [3-5]. Detekcja napadów nie stanowi już problemu, wyzwaniem natomiast jest predykcja napadów. W pracy [6] porównywano skuteczność predykcji napadów metodami liniowymi i nieliniowymi. Warto zauważyć, że chociaż większość danych EEG jest obecnie cyfrowa i dostępnych

¹ p.lichograj@dyd.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl

² m.chodyka@dyd.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl

³ v.golovko@dyd.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl

⁴ bujnik22048@stud.akademiabialska.pl, Zakład Informatyki, Wydział Nauk Technicznych, Akademia Bialska Nauk Stosowanych im. Jana Pawła II, www.akademiabialska.pl

jest wiele protokołów automatycznego wykrywania napadów [7-12], sygnał EEG jest nadal w dużej mierze analizowany za pomocą oględzin.

Zaproponowany w pracy nowy zautomatyzowany algorytm służy do wykrywania napadów u pacjentów z padaczką. Autorska metoda automatycznego wykrywania napadów padaczkowych i wyładowań międznapadowych została opracowana w oparciu o maksymalny krótkoterminowy wykładnik Lyapunov'a (ang. *Lyapunov exponent*) oraz miarę dynamicznej niestabilności układu, która była szeroko stosowana w analizie EEG [7, 8]. Podstawową ideą tej pracy jest dokładne wykrywanie segmentów EEG o różnym czasie trwania z aktywnością padaczkową i niepadaczkową. Pozwala to na identyfikację aktywności patologicznej w stanie remisji oraz wykrycie elementów napadowych w czynności przednapadowej. Zaproponowana technika z wykorzystaniem sieci neuronowych ma zastosowanie zarówno do segmentacji czasowej sygnałów EEG, jak i do obliczania krótkoterminowej maksymalnej wartości wykładnika Lyapunov'a – (STLmax, ang. *Short-term largest Lyapunov exponent*). Napad padaczkowy charakteryzuje się spadkiem STLmax i autorzy proponują wykorzystanie zmiany STLmax w czasie jako kryterium napadu padaczkowego w segmentach EEG. Przyjęty system diagnostyczny generuje dwuwymiarową mapę, która może wizualizować strefy (segmenty) aktywności padaczkowej i niepadaczkowej w sygnałach EEG. Techniki sieci neuronowych pozwalają skrócić czas diagnostyki i liczbę błędnych diagnoz, a także pomóc lekarzowi w podjęciu decyzji. Badania wykonano na danych z Miejskiego Szpitala w Mińsku (Białoruś), jak również wykorzystano publicznie dostępne dane [13]. Skuteczność wykrywania napadów padaczkowych ilustrują wyniki eksperymentalne.

2. Charakterystyka problemu

Badania EEG to nieinwazyjna metoda diagnostyczna pozwalająca zbadać bioelektryczną czynność mózgu. Wykonuje się je w celu lokalizacji ogniska padaczkowego, predykcji wystąpienia napadu, jak również zbadania ewolucji czasowej napadu w celach diagnostycznych. EEG to metoda wykorzystująca sygnały elektryczne do rejestrowania aktywności mózgu. Jest to suma potencjałów postsynaptycznych wielu neuronów w korze mózgowej w postaci wielokanałowego zapisu reakcji w ośrodkowym i autonomicznym układzie nerwowym [14]. Centralny układ nerwowy, w tym mózg i rdzeń kręgowy, reaguje świadomie na bodźce zewnętrzne. Autonomiczny układ nerwowy to system kontroli, który działa nieświadomie, aby regulować funkcje organizmu, takie jak częstota akcji serca, oddychanie i reakcja źrenic. Sygnały EEG odzwierciedlają czynności neuronalne generowane przez bodźce wewnętrzne i zewnętrzne, a także reprezentują informacje fizjologiczne i behawioralne. Badanie polega na umieszczeniu na głowie elektrod i rejestrowaniu aktywności mózgu, zmian potencjału elektrycznego na powierzchni skóry, przy pomocy specjalnego urządzenia – encefalografu.

Analiza stanów epileptycznych może być prowadzona za pomocą liniowych i nieliniowych metod [2]. Istnieje wyraźna różnica we właściwościach dynamicznych sygnałów EEG w stanie niepadaczkowym i padaczkowym. Napady padaczkowe charakteryzują się zsynchronizowanym odpaleniem neuronów, co zmniejsza złożoność zapisu EEG. STLmax zmniejsza się podczas aktywności padaczkowej, dlatego też zaproponowano obliczenie STLmax jako składnik protokołów wykrywania napadów [7, 8, 10]. Jednak konwencjonalne podejścia do obliczania wykładnika STLmax są bardzo wrażliwe na ilość danych i wymagają dużej mocy obliczeniowej [11]. W celu oszacowania

STLmax dla danych EEG w [7] zastosowano zmodyfikowany algorytm Wolfa [8]. Istniejące podejścia mają jednak następujące wady: są niewiarygodne w przypadku małych zbiorów danych i wymagają dużej mocy obliczeniowej. Dlatego wielu autorów używa do obliczania długich segmentów EEG STLmax o długości czasu 10,24 s [15]. Jednym z ograniczeń dotychczasowych badań jest to, że nie pozwalają one dokładnie wykryć małych segmentów EEG z aktywnością padaczkową i niepadaczkową.

Wykrywanie napadów za pomocą protokołów uczenia maszynowego często odbywa się w dwóch etapach: selekcji cech i klasyfikacji zdarzeń [10, 16]. Jednak obecne podejścia do selekcji cech mają te wady, że nie można wybrać segmentów czasowych z napadem padaczkowym, a wymóg uczenia sieci neuronowej na pożądanym zbiorze danych wymaga pewnej ilości niezautomatyzowanej analizy EEG w celu zidentyfikowania reprezentatywnych epileptycznych i niepadaczkowych segmentów EEG, które mają być użyte jako szablony dla zautomatyzowanego algorytmu [1, 7, 8, 10].

3. Materiały i metody

3.1. Baza danych

Do testowania proponowanego podejścia wykorzystano dwa zestawy danych. Pierwszy opisany w [17] i publicznie dostępny w [13]. Kompletny zbiór danych [17] obejmuje pięć podzbiorów (oznaczenia A, B, C, D i E), z których każdy zawiera 100 jednokanałowych sygnałów EEG o czasie trwania 23,6 s i częstotliwości próbkowania 173,6 Hz. Zestaw A i B zostały pobrane z powierzchniowych zapisów EEG pięciu zdrowych ochotników z oczami otwartymi (zestaw A) i zamkniętymi (zestaw B), przy użyciu międzynarodowego schematu rozmieszczenia elektrod 10-20. Zestaw C i D składa się z nagrań od pięciu pacjentów z padaczką. Dane pobrano w okresach beznapadowych. Zestaw E obejmuje aktywność napadową, wybraną ze wszystkich rejestrowanych miejsc wykazujących aktywność napadową.

Kolejna baza danych została zebrana od ośmiu dorosłych pacjentów z aktywnością padaczkową. Badanie pacjentów prowadzono z wykorzystaniem 16-kanałowej rejestracji EEG. Sygnały EEG rejestrowane były z częstotliwością próbkowania 250 próbek na sekundę. Czas trwania jednej rejestracji wynosił około 30 minut. W wyniku przetworzenia tych danych powstała baza danych EEG, która reprezentuje zbiór 21 rejestracji 16-kanałowego EEG, wybranych spośród ośmiu dorosłych pacjentów z padaczką w ciągu 8 sekund dla każdej rejestracji. Można zauważyć, że każdy sygnał w EEG został przedstawiony jako szereg czasowy 2000 punktów. Należy zauważyć, że praktycznie niemożliwe jest wskazanie zdarzenia napadowego w tych sygnałach EEG, nawet dla wysokiej jakości ekspertów neurologów. Celem prowadzonego badania było wykrycie w sygnałach EEG segmentów z aktywnością padaczkową i niepadaczkową.

3.2. Największy wykładnik Lyapunov'a

Wykładnik Lyapunov'a jest statystyczną miarą rozbieżności między dwiema orbitami [18-20]. Niech d_0 będzie początkową rozbieżnością między dwiema trajektoriami, a d_n będzie rozbieżnością między takimi orbitami po n krokach. Wtedy największy wykładnik Lyapunov'a jest określony przez wzór 1.

$$\lambda = \lim_{n \rightarrow \infty} \left(\frac{1}{n} \ln \frac{d_n}{d_0} \right) \quad (1)$$

Konwencjonalne podejście do obliczania STLmax może być stosowane na danych eksperymentalnych tylko wtedy, gdy sekwencja danych jest bardzo długa [8, 11]. Zatem zgodnie z [18] wielkość zbioru danych powinna wynosić $N > 10D$. Jednak spełnienie warunku stacjonarności jest niekiedy trudne i zakłada się *quasi*-stacjonarność dla pewnych odcinków. Dlatego zastosowanie tradycyjnego podejścia do wielu danych w świecie rzeczywistym zostało ograniczone. Jednym ze sposobów uniknięcia tego problemu jest wykorzystanie sieci neuronowych do obliczania STLmax. Metoda opiera się na technikach prognozowania sieci neuronowych do pomiaru dywergencji wykładniczej i pozwala dobrze pracować z małymi zbiorami danych. Sieć neuronową pozwala śledzić trajektorię dowolnego punktu w zrekonstruowanej przestrzeni fazowej, a po zbudowaniu z odpowiednim parametrem zanurzenia (wymiarem zanurzenia D i opóźnieniem czasowym τ) pozwala na dokładniejszą predykcję szeregu czasowego oraz rekonstrukcję atraktora z dowolnego punktu początkowego. W rezultacie prognozująca sieć neuronowa może zachować dynamikę systemu. Oznacza to, że dla każdego punktu w atraktorze można wziąć punkt najbliższy oddalony od niego o pewną odległość, a następnie można prześledzić jego trajektorię.

3.3. Opis systemu

W tej części opisano system diagnostyczny sieci neuronowej do wykrywania napadów padaczkowych z wykorzystaniem danych EEG. Jako kryterium diagnostyczne przyjmuje się wartość STLmax, która ulega obniżeniu w czasie napadów padaczkowych. STLmax charakteryzuje wrażliwość na warunki początkowe [8, 11]. System diagnostyczny sieci neuronowej składa się z różnych modułów, które łączą się w system diagnostyczny (rys. 1).

Dane wejściowe systemu to wielokanałowe dane pacjenta EEG. Dane te można interpretować jako obserwację chaotycznego układu dynamicznego generującego fale elektrofizjologiczne. Dane EEG zarejestrowane z elektrod skóry głowy zawierają różne artefakty i składają się z różnych kombinacji sygnałów. Dlatego też w pierwszym etapie do usuwania artefaktów wykorzystuje się analizę składowych niezależnych (ICA, ang. *Independent Component Analysis*).

W drugim etapie każdy sygnał EEG jest dzielony na *quasi*-stacjonarne segmenty przy użyciu algorytmu segmentacji. Segment nazywany jest *quasi*-stacjonarnym, gdy jego zachowanie nie zmienia się pod wpływem przesunięcia czasowego. Wielowarstwowe sieci neuronowe (MLP, ang. *Multi-Layered Perceptron*) służą do adaptacyjnej segmentacji sygnału EEG. Minimalna początkowa długość odcinka EEG wynosi 70 punktów i zmienia się podczas segmentacji adaptacyjnej. Obliczenie STLmax dla każdego wyodrębnionego segmentu jest wykonywane na trzecim etapie. W rezultacie uzyskuje się sekwencję STLmax dla każdego sygnału EEG:

$$\lambda(t) = (\lambda_1, \lambda_2, \dots, \lambda_p) \quad (2)$$

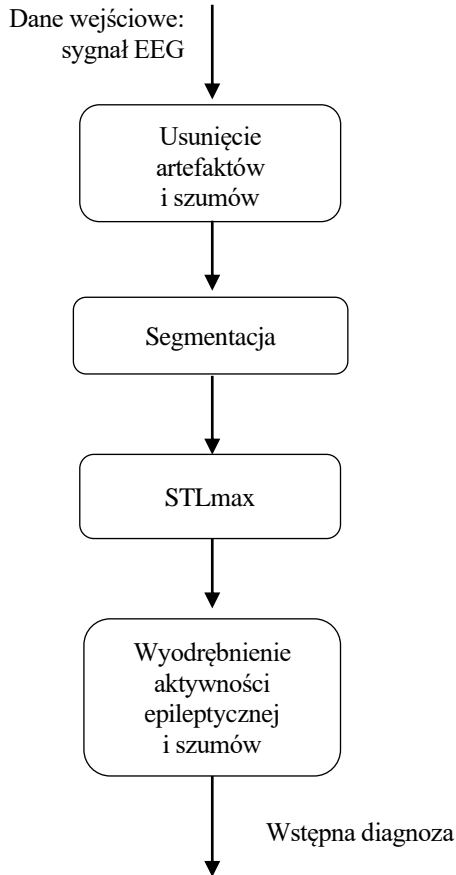
gdzie λ to wykładnik Lyapunov'a, p jest liczbą wybranych segmentów.

W rezultacie otrzymuje się sekwencje segmentów o różnych wartościach STLmax. Jeśli różne segmenty mają taką samą wartość STLmax, są one łączone w jeden segment.

Ostatecznie identyfikację napadu padaczkowego przeprowadza się zgodnie z następującym testem:

$$\begin{cases} \lambda > 0, \text{normalna aktywność} \\ \lambda \leq 0, \text{napad padaczkowy} \end{cases} \quad (3)$$

W efekcie uzyskuje się segmenty w sygnałach EEG o aktywności epileptycznej i normalnej.



Rysunek 1. Schemat blokowy metody do wykrywania aktywności epileptycznej. Zbiór sygnałów EEG jest wykorzystywany jako dane wejściowe do systemu. $\lambda(t)$ to szereg czasowy o wartości STLmax.

ICA – analiza składowych niezależnych [opracowanie własne]

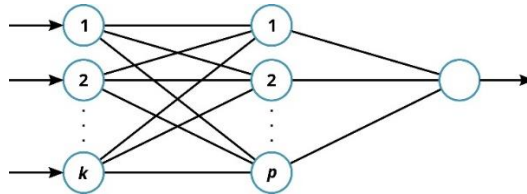
3.4. Segmentacja

Segmentacja to podział sygnału EEG na *quasi*-stacjonarne strefy (segmenty), w których zachowanie nie zmienia się wraz z przesunięciem czasowym. Zaproponowany adaptacyjny algorytm segmentacji służy do dekompozycji sygnału EEG na elementarne interwały. Prognozująca sieć neuronowa, czyli perceptron wielowarstwowy

(MLP, ang. *Multilayer Perceptron*) jest wykorzystywana do segmentacji adaptacyjnej. Na początek należy wykonać rekonstrukcję fazowo-przestrzenną jednowymiarowego sygnału EEG. Przestrzeń fazową procesu chaotycznego można zrekonstruować tylko z jednego szeregu czasowego obserwacji przy użyciu parametrów zanurzenia. Rekonstrukcja fazowo-przestrzenna jest odwzorowaniem, w którym każdy punkt $x(t)$ szeregu czasowego wiąże się z punktem $(x(t), x(t + \tau), \dots, x(t + (D - 1)\tau)) \in R^D$, gdzie t – czas skwantowany ($t = 1, N - D\tau$), τ – opóźnienie czasowe, a D – wymiar zanurzenia.

Takens [18] pokazuje, że za pomocą tylko jednowymiarowej obserwacji można skonstruować w przestrzeni pseudofazowej atraktor $(x(t), x(t + \tau), \dots, x(t + (D - 1)\tau))$, który zachowuje podstawowe właściwości topologiczne i dynamiczne atraktora początkowego układu dynamicznego. W tym celu należy przyjąć wymiar zanurzenia jako $D = 2[d] - 1$, gdzie $[d]$ oznacza część całkowitą fraktalnego wymiaru atraktora d , τ jest odpowiednim opóźnieniem czasowym. W tym przypadku można zrekonstruować atraktor w przestrzeni fazowej R^D . Twierdzenie o zanurzeniu mówi, że tylko z jednej miary układu dynamicznego można odtworzyć przestrzeń stanów, równoważną nieznanemu oryginałowi.

Biorąc to pod uwagę, można stwierdzić, że trzeba zbudować sieć neuronową, która ma $k \geq D - 1$ jednostek wejściowych.



Rysunek 2. Przewidywanie sieci neuronowej [opracowanie własne]

Przewidywanie sieci neuronowej obejmuje trzy warstwy. Każda warstwa składa się z k , p i jednego neuronu, zatem algorytm podziału początkowego sygnału EEG na elementarne interwały metodą sieci neuronowej jest następujący:

1. Sygnał EEG jest wstępnie dzielony na krótkie odcinki o długości N punktów ($N = 70$). Punktem początkowym przesuwanego okna jest $t = 1$.
2. Próbkę uczące mają postać: $\{x(t), x(t + 1), \dots, x(t + N - 1)\}$.
3. Perceptron wielowarstwowy jest szkolony za pomocą metody przesuwającego okna.
4. Perceptron zaczyna przewidywać punkty segmentu. W rezultacie otrzymuje się następujące punkty: $\{x'(t + N), x'(t + N + 1), x'(t + N + 2), \dots\}$. Przewidywanie danych kończy się, gdy spełni się wyrażenie (4).

$$|x'(i) - x(i)| > \Delta x_{\max} \quad (4)$$

gdzie $i = t + N, t + N + 1, \dots$, $\Delta x_{\max} = 0,1$ jest odpowiednim błędem prognozy.

5. Jeżeli $i = t + N$ tj. wyrażenie (4) jest spełnione, to i jest punktem granicy segmentu i tworzony jest kolejny zestaw danych uczących, zaczynając od $t = i$. W przeciwnym razie granica segmentu przesuwa się o liczbę przewidywanych punktów, tj. $t = i - N$.
6. Procedura jest kontynuowana, gdy $t < m - N$, gdzie m jest wspólną długością szeregu czasowych.

Po spełnieniu tego algorytmu otrzymuje się zestaw różnych segmentów i wielowarstwowych perceptronów dostosowanych do odpowiednich segmentów.

Sygnal EEG jest dzielony na początkowe krótkie *quasi*-stacjonarne segmenty. Jakość segmentacji zależy od początkowej długości segmentów [22]. Wybrana została początkowa długość odcinka równa 70 punktom, co daje w przybliżeniu 0,3 s. 70 punktów to wystarczająca długość segmentu do poprawnego trenowania sieci neuronowej w zadaniu różnych zdarzeń epileptycznych. Gdy $N < 70$, liczba błędów fałszywie dodatnich jest zwiększona, ponieważ nie ma wystarczającej liczby punktów, aby wytrenować sieć neuronową do jakościowego prognozowania. Kiedy $N > 70$, prawdopodobieństwo wykrycia pojedynczego krótkiego epizodu padaczkowego jest zmniejszona. Taka długość segmentów początkowych została potwierdzona eksperymentalnie, a zachowanie EEG wewnątrz każdego segmentu jest stałe.

Następnie perceptron wielowarstwowy jest szkolony dla pierwszego segmentu metodą przesuwającego okna:

$$x(t + i\tau) = F(x(t + (i-1)\tau), x(t + (i-2)\tau), \dots, x(t + (i-k)\tau)) \quad (5)$$

gdzie $i = \overline{k, n}$.

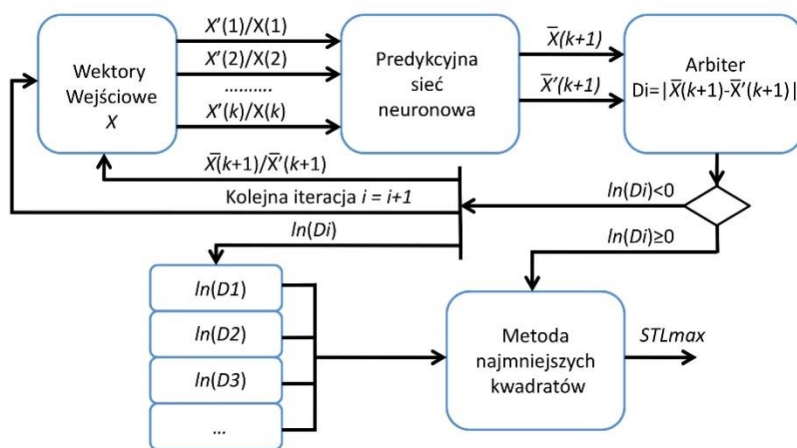
We wzorze τ jest opóźnieniem czasowym, które jest obliczane jako pierwsze minimum metody wzajemnej informacji [18], k jest rozmiarem przesuwającego okna. Należy zauważyć, że $k \geq D - 1$, gdzie D jest wymiarem zanurzeniowym. Wymiar zanurzenia oblicza się metodą fałszywych najbliższych sąsiadów [20]. Eksperymenty pokazują, że wymiar zanurzenia zmienia się od 7 do 9. Perceptron wielowarstwowy składa się z 7 jednostek warstwy wejściowej, 5 jednostek warstwy ukrytej i 1 jednostki warstwy wyjściowej (rys. 2). Po wytrenowaniu MLP granica pierwszego segmentu jest określana za pomocą dokładności prognozowania. Zatem dokładność prognozy jest kryterium wykrywania granicy. Następnie wybieranych jest kolejnych 70 punktów w sygnale EEG, zaczynając od granicy pierwszego segmentu, i z wykorzystaniem punktów tego segmentu trenowany jest drugi MLP. Granica drugiego segmentu jest określana przy użyciu dokładności przewidywania.

Zatem kluczową ideą proponowanej metody jest obliczenie za pomocą dywergencji sieci neuronowej między dwiema orbitami na n kroków do przodu, przy użyciu podejścia iteracyjnego. W takim przypadku procedurę obliczania STLmax przy użyciu małego zbioru danych można opisać zgodnie z następującym algorytmem [19]:

1. Trenuj sieć neuronową za pomocą techniki przesuwającego okna.
2. Wybierz pierwszy punkt $x(t)$, $t = 1$ z wybranego odcinka i utwórz następujący punkt danych: $\{x(t), x(t + \tau), \dots, x(t + (k-1)\tau)\}$, gdzie k to rozmiar okna.
3. Oblicz $\{x(t + k\tau), x(t + (k+1)\tau), \dots, x(t + n\tau)\}$, używając predykcji wieloetażowej.
4. $x(t + i\tau) = F(x(t + (i-1)\tau), x(t + (i-2)\tau), \dots, x(t + (i-k)\tau))$, gdzie $i = \overline{k, n}$, F jest funkcją nieliniową.
5. Oblicz $x'(t) = x(t) + d_0$, gdzie $d_0 \approx 10^{-8}$ i używaj nowego wzorca wprowadzania $\{x'(t), x(t + \tau), \dots, x(t + (k-1)\tau)\}$ powtórz krok 3, aby otrzymać $x'(t + i\tau)$, $i = \overline{k, n}$.

6. Zdefiniuj $\ln d_i = \ln|x'(t + i\tau) - x(t + i\tau)|$, $i = \overline{k, n}$ i zaznacz punkt, dla którego $\ln d_i < 0$.
7. Sporządź wykres $\ln d$ w funkcji n .
8. Zbuduj linię regresji dla zaznaczonego punktu i oblicz jej nachylenie, które jest równe STLmax dla $t = 1$.
9. Zmień czas $t = t + 1$ i powtórz ten algorytm od kroku 3 dla wszystkich punktów odcinka.
10. Oblicz STLmax jako średnią arytmetyczną wszystkich STLmax.

Na rysunku 3 przedstawiono procedurę obliczania STLmax dla przypadku $t = 1$. Wektor wejściowy X składa się z punktów $X(1)$, $X(2)$ i tak dalej. Najbliższa trajektoria dla X to $X' = \{X'(1), X(2) \dots X(k)\}$, k to liczba neuronów wejściowych sieci neuronowej, jest wynikiem prognozowania sieci neuronowej, D_i to odległość między prognozami dwóch najbliższych trajektorii.



Rysunek 3. Schemat obliczeń STLmax [opracowanie własne]

Dzięki zastosowaniu tej techniki STLmax dla szeregów czasowych Henona i Lorenza wynoszą odpowiednio 0,43 (pożądana wartość to 0,418) i 0,98 (pożądana wartość to 0,906). W obu przypadkach zastosowano tylko serię X . Wielkość zbioru danych wynosiła odpowiednio 70 i 100 punktów. Jak widać, wynik ten jest bardzo zbliżony do wartości rzeczywistych z błędem odpowiednio 2,87% i 8,16%. Wyniki są lepsze w porównaniu z [11], gdzie błąd wynosi 16,7% dla 1000 punktów szeregów czasowych Lorenza i 1,9% dla 100 punktów szeregów czasowych Henona. Ta metoda jest bardzo korzystna ze względu na złożoność obliczeniową, dokładność i mały zbiór danych.

Należy zauważyć, że do obliczenia STLmax wykorzystane zostały wielowarstwowe sieci neuronowe, otrzymane na etapie segmentacji. Jeśli różne segmenty mają taką samą wartość wykładnika STLmax, są one łączone w jeden segment. Na koniec przeprowadza się identyfikację napadu padaczkowego dla każdego segmentu. Zaproponowane podejście pozwala na dokładne wykrycie epileptycznych i niepadaczkowych segmentów EEG o różnym czasie trwania i kształcie, w celu identyfikacji aktywności patologicznej w stanie remisji oraz wykrycia aktywności napadowej w okresie przednapadowym.

4. Wyniki doświadczeń

W badaniach wykorzystane zostały zestawy sygnałów EEG. W każdym zestawie znajduje się 100 segmentów EEG. Każdy segment EEG zawiera 4096 kolejnych punktów amplitudy, jego czas trwania wynosi 23,6 sekundy przy próbkowaniu 173,61 Hz. Przeprowadzone zostały eksperymenty na sygnałach EEG charakteryzujących aktywność epileptyczną i normalną. Sygnały podzielono na dwie klasy: klasa 1 obejmowała sygnały z aktywnością padaczkową wykrytą za pomocą proponowanego systemu, klasa 2 obejmowała sygnały tylko z aktywnością niepadaczkową.

W tabeli 1 przedstawiono wyniki klasyfikacji dla zestawów AE. Każdy zestaw składa się ze 100 sygnałów EEG.

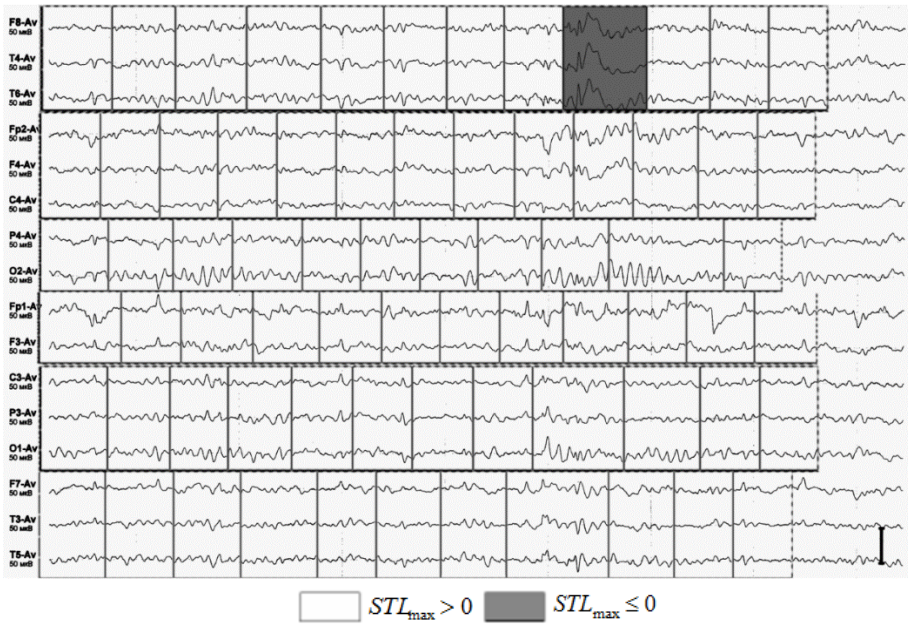
Tabela 1. Wyniki klasyfikacji dla zestawów AE

Zestawy	Ilość segmentów z napadami [%]	Ilość segmentów bez napadów [%]
A	0	100
B	0	100
C	6	94
D	32	68
E	92	8

Źródło: opracowanie własne.

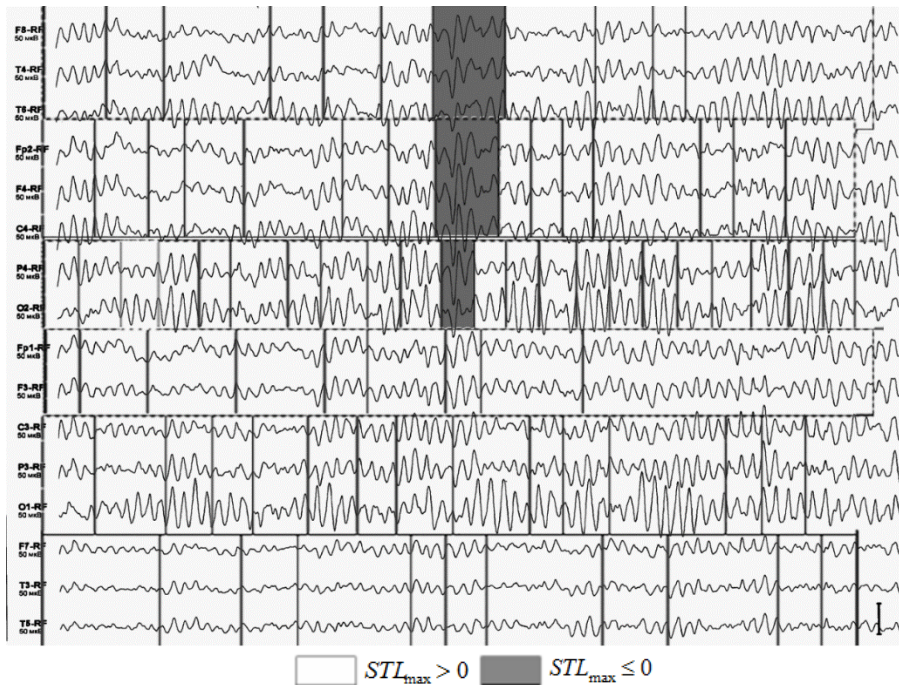
Znamienne jest, że nie ma fałszywych detekcji aktywności padaczkowej w zestawach A i B. W leczeniu chirurgicznym konieczne jest znalezienie strefy padaczkowej (źródła napadów padaczkowych). Co ciekawe, 6% detekcji aktywności padaczkowej w zbiorze C to tylko pojedyncze detekcje (wykrywany jest tylko jeden segment z aktywnością padaczkową w każdym sygnale). Gdy analizowano zestawy C i E, w większości przypadków system wykrył wiele detekcji aktywności padaczkowej. Wyniki pokazały, że system może nie tylko wykryć obecność aktywności epileptycznej, ale także przydzielić segmenty, w których się ona znajduje.

W kolejnym eksperymencie wykorzystano dane EEG dostarczone przez Miejski Szpital w Mińsku (Białoruś). Dane reprezentują zestaw 21 rejestracji 16-kanalowego EEG. Dane EEG rejestrowano od ośmiu dorosłych pacjentów przez 8 sekund dla każdej rejestracji. W wyniku digitalizacji danych EEG z częstotliwością 250 Hz każdy sygnał w EEG został przedstawiony jako szeregi czasowe po 2000 punktów. Łączna liczba sygnałów EEG wynosi 336. W porównaniu z poprzednią bazą, dane te zawierają różne artefakty. Jak wspomniano wcześniej, do usuwania artefaktów z zapisów EEG wykorzystana została analiza składowych niezależnych (ICA). Sygnały EEG jednej rejestracji są podzielone na sześć zestawów. Wszystkie zapisy jednej rejestracji są podzielone na sześć zestawów. Ilość zestawów dobierana jest doświadczalnie w zależności od ilości sygnałów w zestawie dla właściwego filtra oraz rozmieszczenia elektrod. Następnie każdy zestaw sygnału EEG jest przetwarzany przez moduł ICA. W efekcie otrzymano sześć pozbawionych artefaktów sygnałów EEG. Następnie przeprowadzana jest segmentacja każdego uzyskanego sygnału EEG. W rezultacie wykryto 1775 segmentów z sygnałów EEG wszystkich rejestracji ($6 \times 21 = 126$). Wyniki wykrywania padaczki dla jednej rejestracji przedstawiono na wykresach, jak pokazano na rysunkach 4, 5 i 6.



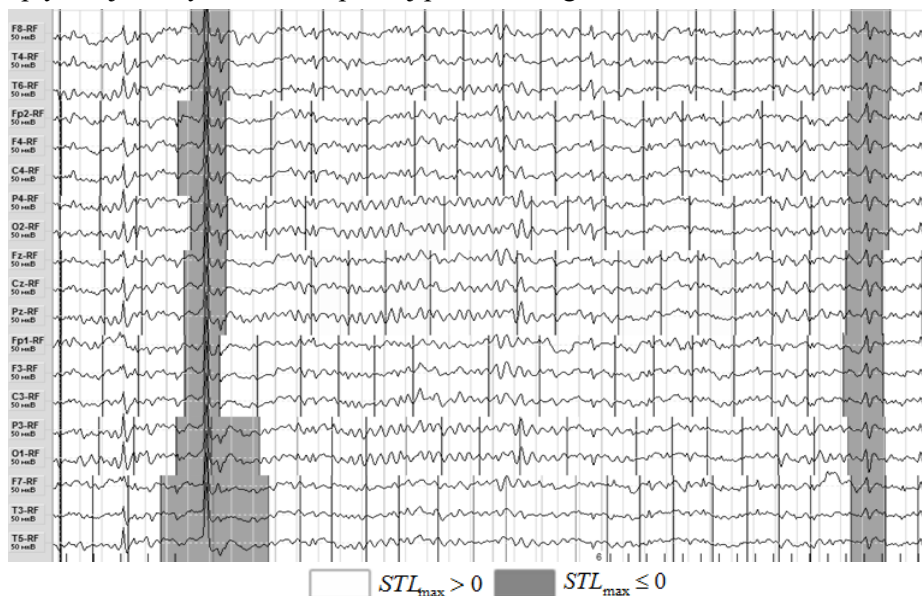
Rysunek 4. Wyniki detekcji aktywności padaczkowej dla zestawu sygnałów EEG [opracowanie własne]

Na rysunku 4 system wykrywa jeden segment w części skroniowej prawej półkuli mózgu (zaznaczone szarym kolorem).



Rysunek 5. Wyniki detekcji aktywności padaczkowej dla zestawu sygnałów EEG [opracowanie własne]

Kolejny wynik (rys. 5) pokazuje, że system wykrywa jedno ognisko aktywności epileptycznej w całym obszarze prawej półkuli mózgu.



Rysunek 6. Wyniki detekcji aktywności padaczkowej dla zestawu sygnałów EEG [opracowanie własne]

Na rysunku 6 można zauważyć, że system wykrywa dwa ogniska we wszystkich obszarach mózgu. Praktycznie niemożliwe jest wskazanie zdarzenia napadowego w tych sygnałach EEG za pomocą oględzin, nawet dla wysokiej jakości ekspertów neurologów.

Wyniki klasyfikacji wszystkich wybranych segmentów w danych EEG według zaprojektowanego systemu podsumowano w tabeli 2.

Tabela 2. Klasyfikacja aktywności padaczkowej na podstawie szpitalnej bazy danych dla 8 pacjentów

Nazwa	Liczba wszystkich segmentów	Wyniki klasyfikacji	
		Ilość segmentów z napadami	Ilość segmentów bez napadów
Segmanty padaczkowe	32	29	3
Aktywność normalna	1743	4	1739

Źródło: opracowanie własne.

System poprawnie wykrywa 29 segmentów z aktywnością padaczkową z łącznej liczby 32 segmentów. Wydajność testu prezentowanego podejścia można zdefiniować poprzez obliczenie czułości, specyficzności i całkowitej dokładności klasyfikacji.

Zakładając, że FP (fałszywie pozytywne) to liczba normalnych segmentów oznaczonych jako epileptyczne, FN (fałszywie negatywne) to liczba segmentów epileptycznych oznaczonych jako normalne, TP (prawdziwie dodatnie) to liczba poprawnie wykrytych segmentów epileptycznych, TN (prawdziwie ujemne) to liczba prawidłowych segmentów

wykrytych prawidłowo. Następnie metryki oceny obejmują następujące kryteria pomiaru wydajności proponowanego podejścia:

- czułość znana również jako współczynnik prawdziwie dodatnich wyników

$$TPR = Se = \frac{TP}{TP + FN} = \frac{TP}{P}; \quad (6)$$

- specyficzność znana również jako współczynnik prawdziwie ujemny

$$TNR = Sp = \frac{TN}{TN + FP} = \frac{TN}{N}; \quad (7)$$

- całkowita dokładność

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} = \frac{TP + TN}{P + N} \quad (8)$$

Wartości tych parametrów statystycznych obliczone na podstawie tabel 1 (zestawy A, B i E) oraz 2 przedstawiono w tabeli 3.

Tabela 3. Porównanie wydajności na dwóch zbiorach danych

Parametry statystyczne	Wartość dla zestawu A, B, E [%]	Wartość dla danych szpitalnych [%]
Swoistość	100,0	99,7
Czułość	92,0	90,6
Całkowita dokładność klasyfikacji	96,0	99,6

Źródło: opracowanie własne.

Wyniki pokazują, że przedstawiona w pracy metodologia analizy EEG jest bardzo specyficzna (99,7%), co oznacza, że występują niewielkie zliczenia fałszywej detekcji aktywności padaczkowej. To ważne, ponieważ błędna diagnoza może mieć poważne konsekwencje. Wartość czułości oznacza, że system w 90,6% przypadków prawidłowo wykrywa aktywność padaczkową w rzeczywistych danych EEG. Łączna dokładność klasyfikacji segmentów w dwóch klasach (aktywność niepadaczkowa i aktywność padaczkowa) wynosi 99,6%.

5. Podsumowanie

W pracy omówiono metodę wykrywania napadów padaczkowych za pomocą fal EEG. Proponowane podejście opiera się na selekcji różnych segmentów czasowych w sygnałach EEG z aktywnością padaczkową i niepadaczkową. Wartość STLmax służy do klasyfikacji tych segmentów w danych EEG. Techniki sieci neuronowych są proponowane zarówno do segmentacji sygnałów EEG, jak i do obliczania STLmax. Proponowane podejście wykorzystuje te same sieci neuronowe zarówno do wyboru segmentu czasu, jak i do obliczeń STLmax w każdym segmencie przy użyciu małych zbiorów danych. W porównaniu z podejściem konwencjonalnym, pozwala to zarówno na zmniejszenie złożoności obliczeniowej, jak i na ograniczenie czasu obserwacji.

Głównymi zaletami prezentowanej techniki neuronowej jest możliwość selekcji w EEG małych segmentów czasowych z aktywnością padaczkową i normalną, trening

bez pożądanego zbioru danych o aktywności padaczkowej i niepadaczkowej, możliwość pomocy lekarzowi w podjęciu decyzji oraz wizualizacja stref (segmentów) aktywności padaczkowej i niepadaczkowej w sygnałach EEG. Zatem proponowana metoda ma:

- wysoką dokładność klasyfikacji segmentów pod kątem aktywności epileptycznej i niepadaczkowej;
- automatyczne wykrywanie aktywności epileptycznej w EEG;
- możliwość klasyfikacji bez uprzedniego szkolenia na specjalnie pożądanym zbiorze danych;
- możliwość wykrywania aktywności napadowej o różnych kształtach i czasie trwania.

Proponowana metoda pozwala wykryć odcinki EEG o różnym czasie trwania z aktywnością padaczkową i niepadaczkową, wykryć aktywność patologiczną w stanie remisji oraz aktywność napadową w okresie przednapadowym. W związku z tym, iż badanie zostało przeprowadzone na niezbyt dużej bazie danych, dalsze prace badawcze dotyczyć będą większych zbiorów danych, by lepiej ocenić moc generalizacji metody.

Literatura

1. Shankar A., Dandapat S., Barma S., *Classification of Epileptic Seizure From EEG Signal Based on Hilbert Vibration Decomposition and Deep Learning*, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2021.
2. Pijn J.P., Velis D.N., van der Heyden M.J., DeGoede J., van Veelen C.W., Lopes da Silva F.H., *Nonlinear dynamics of epileptic seizures on basis of intracranial EEG recordings*, Brain Topogr, 9(4), 1997, s. 249-270.
3. Aschenbrenner-Scheibe H. et al., *How well can epileptic seizures be predicted? An evaluation of a nonlinear method*, Brain, 126, 2003, s. 2616-2626.
4. Lai Ying-Cheng et al., *Inability of Lyapunov Exponents to Predict Epileptic Seizures*, Phys Rev. Lett, 91(6), 2003, s. 68102.
5. McSharry P.E., Smith L.A., Tarassenko L., *Prediction of epileptic seizures: are nonlinear methods relevant?*, NatMed, 9, 2003, s. 241-242.
6. Morman F., Kreuz T., Rieke Ch., Andrzejak R.G., Kraskov A., David P., Elger Ch.E., Lehnertz K., *On the predictability of epileptic seizures*, ClinNeurophys, 116(3), 2005, s. 569-567.
7. Iasemidis L.D., Principe J.C., Sackellares J.C., *Measurement and quantification of spatiotemporal dynamics of human epileptic seizures*, Nonlinear signal processing in medicine, 2000, s. 1-27.
8. Wolf A., Swift J., Swinney H., Vastano J., *Determining Lyapunov exponents from a time series*, Physica D, 16, 1985, s. 285-292.
9. Sheb A., Gutttag J., *Aslication of machine learning to epileptic seizure detection*, 2010 the 27th International Conference on machinelearning, Haifa, Israel 2010.
10. Derya Ubeyli E., *Statistics over featuers: EEG signals analysis*, ComputBiol Med, (39)8, 2009, s. 733-741.
11. Rosenstein M.T., Collins J.J., De Luca C.J., *A practical method for calculating largest Lyapunov exponents from small data sets*, Physica D, 65, 1993, s. 117-134.
12. Wang X., Meng J., Tan G., Zou T., *Research on the relation of EEG signal chaos characteristics with high-level intelligence activity of human brain*, Nonlinear Biomedical Physics, 2010, s. 4631-4642, doi: 10.1186/1753.
13. EEG time series: <http://meb.unibonn.de/epileptologie/science/physik/eegdata.html>, 2011.
14. Khalifa W., Salem A., Roushdy M., Revett K., *A survey of EEG based user authentication schemes*, 8th International Conference on Informatics and Systems (INFOS), IEEE, BIO-55-BIO-60, 2012.

15. Nair S., Shiau D., Principe J., Iasemidis L., Pardalos P., Norman W., Carney P., Kelly K., Sackellars J., *An investigation of EEG dynamics in an animal model of temporal lobe epilepsy using the maximum Lyapunov exponent*, *Experimental Neurology*, 216, 2009, s. 115-121.
16. Patnaika L.M., Manyam O.K., *Epileptic EEG detection using neural networks and post-classification*, *Computer methods and programs in biomedicine*, 91, 2008, s. 100-109.
17. Andrzejak R., Widman G., Lehnertz K., Rieke C., David P., Elger C., *The epileptic process as nonlinear deterministic dynamics in a stochastic environment: an evaluation on mesial temporal lobe epilepsy*, *Epilepsy Res*, 44, 2001, s. 129-140.
18. Golovko V., Doudkin A., Maniakov N., *Aslication of Neural Networks Techniques to Chaotic Signal Processing*, *Optical Memory and Neural Networks*, 13, 2004, s. 195-215.
19. Golovko V., *From Neural Networks to Intelligent Systems: Selected Aspects of Training, Aslication and Evolution*, In book *Limitations and Future Trends in Neural Computation: NATO book*. IOS Press, Amsterdam 2003, s. 219-243.
20. Golovko V., Savitsky Y., Maniakov N., *Neural Networks for Signal Processing in Measurement analysis and Industrial Aslications: the Case of Chaotic Signal processing*, In book *Neural Networks for Instrumentation, Measurement and related industrial Aslications: NATO book*. IOS Press, Amsterdam 2003, s. 119-144.
21. Takens F., *Detecting strange attractors in turbulence*, *Lecture Notes in Mathematics*, Springer-Verlag, 898, Berlin 1981, s. 366-381.
22. Lopes da Silva F., *EEG analysis: theory and practice*, [w:] Niedermeyer E., Lopes da Silva F., (red.) *Electroencephalography: basic principles, clinical aslications, an related fields*, 4th ed. Baltimore: Lisincott, Williams and Wilkins, 1998, s. 1153-1163.

Metoda wykrywania napadów padaczkowych za pomocą fal EEG

Streszczenie

W ciągu ostatnich kilku dekad zastosowanie sieci neuronowych do analizy elektroencefalogramu (EEG) gwałtownie wzrosło. Przedstawiona metoda wykrywania napadów padaczkowych opiera się na maksymalnym krótkoterminowym wykładniku Lyapunov'a (STLmax). Proponowane podejście opiera się na automatycznej segmentacji sygnałów EEG na części odpowiadające aktywności epileptycznej i niepadaczkowej. STLmax jest następnie obliczany z obu rodzajów sygnału EEG i używany do klasyfikacji epileptycznych i niepadaczkowych segmentów EEG w całym zapisie. Techniki sieci neuronowych są proponowane zarówno do segmentacji sygnałów EEG, jak i do obliczania STLmax. Do przeprowadzenia eksperymentów wykorzystany został zbiór szpitalnych danych złożony z 21 nagrań trwających 8 sekund ośmiu dorosłych pacjentów. Ponadto do eksperymentów wykorzystano publicznie dostępne dane. Proponowana metoda pozwala dokładnie wykryć odcinki EEG o różnym czasie trwania z aktywnością padaczkową i niepadaczkową, aktywność patologiczną w stanie remisji oraz aktywność napadową w okresie przednapadowym.

Słowa kluczowe: perceptron wielowarstwowy, największy wykładnik Lyapunov'a, elektroencefalogram, napad padaczkowy

A method of detecting epileptic seizures using EEG waves

Abstract

Over the past few decades, application of neural networks to electroencephalogram (EEG) analysis has grown rapidly. We report a method for epileptic seizure detection that relies on the maximal short-term Lyapunov exponent (STLmax). The proposed approach is based on automatic segmentation of the EEG into epochs that correspond to epileptic and non-epileptic activity. The STL-max is then computed from both categories of EEG signals and used for classification of epileptic and non-epileptic EEG segments throughout the recording. Neural network techniques are proposed both for segmentation of EEG signals and computation of STLmax. The data set from hospital have been used for experiments performing. It consisted of 21 recordings of EEG of eight adult patients. Furthermore the publicly available data were used for experiments. The proposed method allows to accurately detect EEG episodes of different duration with epileptic and non-epileptic activity, to detect pathological activity in remission and to detect seizure activity in the pre-ictal period.

Keywords: multilayer perceptron, largest Lyapunov exponent, electroencephalogram, epileptic seizure

Badanie zasadności używania kalibracji w pomiarach okulograficznych w celu zastosowania w interfejsie człowiek-komputer

1. Wstęp

Okulograf (ang. *eyetracker*) to urządzenie, które wykorzystuje widmo światła z zakresu bliskiej podczerwieni do śledzenia ruchów gałek ocznych. Rejestracja odbywa się za pomocą umieszczonych w nim kamer. Za pomocą okulografu możliwe jest określenie kierunku patrzenia człowieka poprzez rejestrację chwilowego położenia środka źrenicy [1]. Można wyróżnić kilka rodzajów okulografów, przykładowo rozróżniając ze względu na położenie względem głowy: mobilne (przenośne) oraz nagłowne (przypominające okulary). Naukowcy decydują się na model korzystniejszy dla ich celów naukowych, w zależności ich celów eksperymentu lub innych czynników, np. wieku użytkowników, od ich stopnia poruszania się podczas badania lub lokalizacji tego badania.

Ważnym aspektem okuloigrafii jest możliwość uzyskania pomiaru różnych parametrów podczas badania – do najczęściej wykorzystywanych należą: średnica źrenicy oczu oraz parametry dotyczące ruchu gałek ocznych (fiksacje, sakkady). Fiksacja to moment, w którym gałka oczna znajduje się w fazie bezruchu. W tym czasie pobierana jest informacja wzrokowa z otoczenia i trafia ona do mózgu w celu dalszego jej przetworzenia [2]. Z kolei przejście między jedną fiksacją a drugą nazywa się sakkadą. Cel konkretnego badania sugeruje, jakie parametry należy mierzyć podczas wykonywania eksperymentu, aby móc je następnie efektywnie analizować.

Obecnie okuloigrafia jest szeroko wykorzystywana zarówno w badaniach naukowych, jak i komercyjnych obejmujących wiele zróżnicowanych dziedzin, np. medycynę, edukację, psychologię czy marketing. Dane wynikowe pochodzące od okulografu mogą służyć np. algorytmom jako dane wejściowe wykorzystywane we wspomaganiu wczesnej diagnostyki medycznej. W niniejszym artykule chcemy zwrócić uwagę na jej coraz częstsze zastosowanie w zakresie interakcji człowiek-komputer (ang. *Human Computer Interaction*).

Przykładem w tej dziedzinie są agenty konwersacyjne (ang. *Embodied Conversational Agents, ECA*), którym poświęca się ostatnio wiele uwagi. Jednym z zastosowań jest np. wykorzystanie zachowań niewerbalnych do nawiązania kontaktu z użytkownikiem [3]. Aby dialog pomiędzy człowiekiem a agentem mógł stać się bardziej wiarygodny, agenty te powinny być wyposażone w zdolności komunikacyjne i ekspresyjne nawiązujące do takich, które znamy z kontaktów międzyludzkich (w zakresie mowy, gestów, mimiki twarzy, spojrzenia itp.).

¹ karolina.gabor-siatkowska.dokt@pw.edu.pl, Instytut Telekomunikacji, Wydział Elektroniki i Technik Informatycznych, Politechnika Warszawska.

² Instytut Telekomunikacji, Wydział Elektroniki i Technik Informatycznych, Politechnika Warszawska.

Dzięki zastosowaniu okulografu w badaniach możliwe jest uzyskanie informacji odnoszących się do stanu psychicznego człowieka: np. jego stan emocjonalny [4, 5], obciążenie poznawcze [6] czy poziom koncentracji [7, 8]. Dodatkowo, coraz częściej stosuje się badania okulograficzne jako dodatkowe źródło informacji podczas rozmów użytkowników z agentem dialogowym [9]. W taki sposób uzyskiwane dane są następnie wykorzystywane do wspomagania lub sterowania agentami konwersacyjnymi. Prowadzony dialog człowiek-komputer staje się dzięki temu bardziej naturalny [10], a co za tym idzie, agenty dialogowe są w stanie wejść w bardziej realistyczną rozmowę z ludźmi.

2. Cel przeprowadzonych badań

Celem przeprowadzonych badań jest sprawdzenie, czy podczas korzystania z okulografu firmy Gazepoint GP3 w interfejsie komputer-człowiek można pominąć proces kalibracji. W niniejszych badaniach został wykorzystany okulograf Gazepoint GP3, ponieważ najistotniejszym zagadnieniem było zachowanie jak największego komfortu użytkowników podczas przeprowadzania eksperymentów. Istotnym aspektem jest to, że osoba może zostać badana za pomocą tego okulografu w sposób całkowicie bezdotykowy – w takim sensie rozumiany „nieinwazyjny”. Jest to ważne, ponieważ przeprowadzone badania mają służyć w przyszłości do badań z udziałem pacjentów leczących się psychiatrycznie. Pacjenci będą mogli skorzystać z sesji terapeutycznych poprzez rozmowę z agentem dialogowym wykorzystującym sztuczną inteligencję (agent „Terabot” [10]). Okulograf będzie służył jako dodatkowe źródło informacji o dialogu pomiędzy pacjentem a Terabotem, w celu ulepszenia jego działania. Głównym celem autorów jest zbadanie, czy kalibracja urządzenia będzie mogła być pominięta, aby w żaden sposób nie naruszać komfortu pacjenta ani nie zmuszać go do długich procedur. Możliwość pominięcia kalibracji urządzenia bardzo korzystnie wpłynęłaby na przebieg sesji terapeutycznych.

3. Metodyka wykonanych badań

3.1. Stanowisko

Na Politechnice Warszawskiej w jednym z laboratoriów Wydziału Elektroniki i Technik Informacyjnych znajdowało się stanowisko składające się z:

- laptopa 15,3'-calowego (procesor IntelCore i7, karta graficzna Nvidia GeForce GT 650M);
- podłączonego do niego okulografu przenośnego, wolnostojącego firmy Gazepoint GP3 wraz z zainstalowanym oprogramowaniem Gazepoint Control oraz Gazepoint Analysis.

Jak już wcześniej wspomniano, okulografy wolnostojące nie zaburzają dialogu z użytkownikami czy sesji terapeutycznej z pacjentami, tak jak robią to okulografy nąglowne. Podczas rozmowy osoba badana może się dość swobodnie poruszać, jej ruch nie jest ograniczony przez dodatkowe kable czy uczucie dyskomfortu, jakie w pewnym stopniu występuje podczas noszenia nakładanych okulografów.

Oświetlenie podczas eksperymentu było naturalne (światło dzienne), równomiernie rozproszone, wpadające przez okno, dopasowane za pomocą żaluzji.

Każdy uczestnik badania siedział na regulowanym fotelu, jego ruch nie był skrępowany, dbano o możliwie największą swobodę i komfort podczas przeprowadzenia ba-

dania. Uczestnik siedzący na fotelu był przysunięty niedaleko biurka, na którym mieścił się laptop oraz okuloGRAF. Odległość pomiędzy głową osoby badanej a monitorem laptopa i okuloGRAfem wynosiła ok. 40-50 cm. OkuloGRAF został umieszczony bezpośrednio pod monitorem, za klawiaturą w taki sposób, aby możliwa była detekcja twarzy uczestnika ze szczególnym uwzględnieniem oczu. Zastosowane urządzenie ma częstotliwość próbkowania 150 Hz, jego dokładność wynosi $0,5-1^\circ$, jak podaje producent.

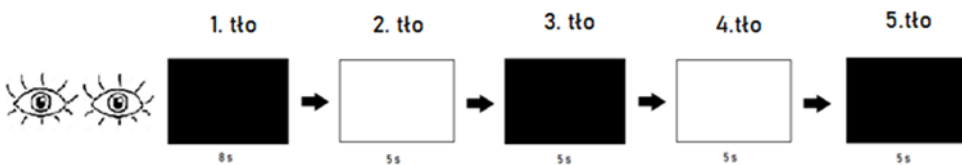
3.2. Grupa badana

Niniejsze eksperymenty zostały przeprowadzone z udziałem 32 osób: 16 kobiet oraz 16 mężczyzn (w przedziale wiekowym od 21 do 30 lat), studentów i studentek Politechniki Warszawskiej i Uniwersytetu Warszawskiego.

Wśród tych osób 13 miało założone soczewki lub okulary korekcyjne. Podczas badań celowo nie prosiło uczestników o zdjęcie okularów, mimo że mają one wpływ na rejestrację danych pochodzących z okuloGRAfu. Autorzy podjęli świadomie to ryzyko, mając na uwadze specyficzne warunki, jakie panują np. w przychodniach czy szpitalach. Przy takiej specyfice badań nie można prosić pacjentów o zdjęcie soczewek czy okularów, ponieważ miałyby to negatywny wpływ na sesje terapeutyczne oraz w efekcie uczyniłyby je bezsensownymi. Pacjenci psychiatryczni mogliby odczuć dyskomfort podczas takiej sesji, mogłoby to również doprowadzić do pogorszenia ich stanu emocjonalnego/psychicznego. Autorzy uważają, że nawet zniekształcone dane uzyskane poprzez pozostawienie pacjentowi okularów czy soczewek są nadal lepsze niż ich całkowity brak. Dzięki takiemu podejściu warunki eksperymentu w laboratorium były zbliżone do warunków przeprowadzania badań podczas sesji terapeutycznych u pacjentów psychiatrycznych.

3.3. Opis przeprowadzonych eksperymentów

Badanie każdego uczestnika polegało na oglądaniu ekranu laptopa, podczas którego wyświetlane były następujące po sobie całkowicie czarne i całkowicie białe tła. W tym czasie pobierane były dane pochodzące z okuloGRAfu na temat zachowania się źrenic oraz ruchu gałek ocznych. Czas trwania wyświetlania każdego tła wynosił 5 s z wyjątkiem pierwszego czarnego tła, które ukazywało się przez 8 s. Ten dodatkowy czas przy wyświetlaniu pierwszego tła służył akomodacji oczu, a konkretnie źrenic do zaistniałych warunków. W sumie czas wyświetlania całości wynosił 28 s. W tym czasie okuloGRAF monitorował ruch gałek ocznych uczestnika, rozpoznawał jego prawą i lewą źrenicę. Każdy uczestnik oglądał zmieniające się tła dwukrotnie (z zachowaniem odpowiedniej przerwy pomiędzy tymi podejściami): pierwszy raz oglądanie i pobieranie danych nastąpiło bez wcześniejszej kalibracji urządzenia, natomiast za drugim razem przeprowadzana była kalibracja zapewniana przez oprogramowanie firmy Gazepoint. Rysunek 1 pokazuje schematycznie przebieg całego badania.



Rysunek 1. Schemat zmieniającego się tła podczas badania z okuloGRAfem [opracowanie własne]

W przypadku eksperymentu bez kalibracji dbano tylko o odpowiednią odległość oczu osoby badanej do okulografu (przy pomocy oprogramowania Gazepoint Control). Stała detekcja oczu i źrenic w tym programie umożliwiła ustalenie odpowiedniej odległości umieszczenia głowy i oczu osoby badanej do samego okulografu. Program dawał również informację zwrotną o odległości oczu od urządzenia (za blisko/optymalnie/za daleko). Przy pozytywnej odpowiedzi można było przejść do wykonania badania.

W przypadku eksperymentu z kalibracją badanie było poprzedzone procedurą kalibracji zapewnioną przez dostępne oprogramowanie. Istotne było, aby badania przebiegły w warunkach komfortowych dla użytkownika oraz aby były przeprowadzone w możliwie najbardziej zbliżony sposób do warunków podczas sesji terapeutycznych. Uczestnik mógł swobodnie mrugać, ruszać głową, nie było żadnych ograniczających reguł. W praktyce okazało się, że wszyscy uczestnicy zachowywali się naturalnie; osoby siedziały wyprostowane i spokojnie patrzyły się na ekran laptopa, nie wykonując żadnych nadmiernych ruchów.

3.4. Rejestrowane dane

W niniejszym eksperymencie do analizy użyto danych od wszystkich 32 uczestników. Podczas badania z kalibracją oraz bez niej rejestrowano szereg parametrów okulograficznych. Do analizy wykorzystano następujące z nich:

- wielkość średnicy prawej i lewej [mm];
- czas trwania fiksacji [s];
- czas trwania mrugnięcia [s].

W literaturze opisany jest szereg różnych parametrów, które mają wpływ na zachowanie się średnicy źrenicy człowieka. Podstawową funkcją źrenicy jest dostosowywanie wpadającego światła do wnętrza oka poprzez jej zmianę wielkości, co oznacza, że źrenica reaguje na zmianę intensywności światła. Poza tym reaguje ona również m.in. na:

- stan koncentracji człowieka [8, 11];
- obciążenie poznawcze [6, 12];
- usłyszane dźwięki (przeżywanie pozytywnych czy negatywnych emocji) [13];
- oglądane obrazy (przeżywanie pozytywnych czy negatywnych emocji) [14];
- spożywanie alkoholu, narkotyków itp. [15].

Tak jak wypisano powyżej, wiele czynników może mieć wpływ na średnicę źrenicy oka. Jednak podczas wykonywania badania u studentów i studentek, autorzy zapewnili możliwie jednorodny warunki podczas eksperymentu w następujący sposób:

1. Obciążenie poznawcze: Badania zostały wykonane u wszystkich uczestników po przeprowadzeniu zajęć laboratoryjnych, nie były to również pierwsze zajęcia dla tej grupy osób w danym dniu. Wówczas, obojętnie czy dany temat omawiany był dla kogoś łatwiejszy czy trudniejszy, laboratorium zostało zakończone i po kilkuminutowej przerwie każdy z uczestników deklarował, czy ma ochotę wziąć udział w badaniu. Kilka osób nie miało czasu lub czuły się zmęczone – nie były brane w ogóle pod uwagę w próbie badawczej.
2. Samopoczucie/stan zdrowia: Autorzy przeprowadzili krótkie rozeznanie w postaci zapytania „Jak się czujesz?”. Odpowiedzi osób były pozytywne/neutralne („dobrze” lub „ok”). Wśród wszystkich uczestników laboratoriów zdarzyły się osoby, które nie potwierdziły dobrego samopoczucia, zostały wówczas wyłączone z badania.

3. Alkohol, używki, narkotyki itp.: Ten efekt został wyeliminowany poprzez przeprowadzenie badania po odbyciu zajęć laboratoryjnych – żadna z powyższych osób nie wykazywała odmiennego zachowania, które mogłoby być skutkiem zażywania substancji psychoaktywnych.
4. Usłyszane dźwięki/oglądane obrazy – w laboratorium panowała cisza, nie było żadnego zakłócenia w postaci nagłego huku/wrzasków/syren pojazdów itp. Ponadto wszyscy studenci przebywali w tym samym pomieszczeniu kilka minut przed badaniem, dzięki czemu ich oczy mogły się dostosować do zaistniałych warunków wizualnych i oświetleniowych.
5. Emocje pozytywne/negatywne – podczas pracy na laboratorium panowała spokojna atmosfera, z kolei podczas kilkuminutowej przerwy bezpośrednio przed badaniem proszono wszystkich o zachowanie ciszy, dlatego można wywnioskować, że takie wyciszenie miało u wszystkich taki sam efekt i emocje nie miały wpływu na badanie. Dodatkowo wybrano celowo do eksperymentu tło czarne oraz tło białe, czyli neutralne kolory ekranu, które w żaden sposób nie wpływałyby na stan emocjonalny studentów.

4. Analiza i wnioski

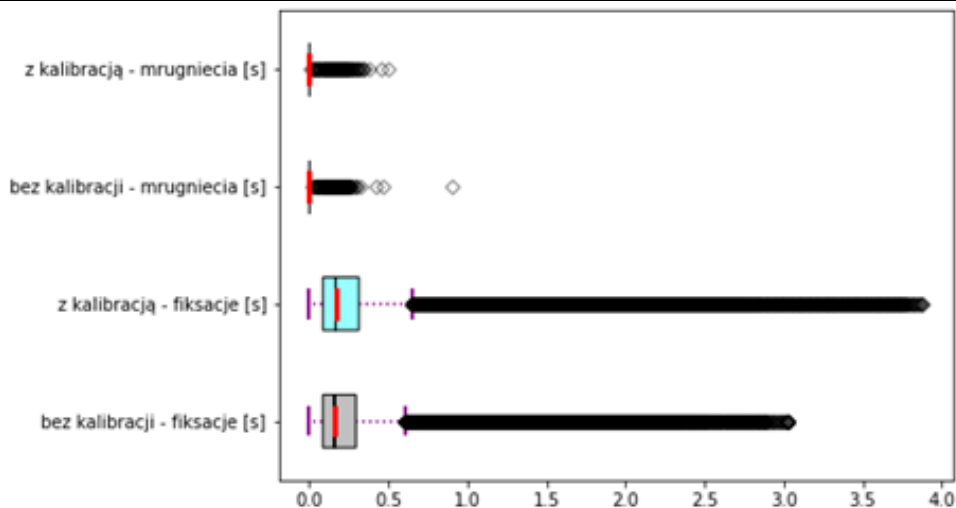
4.1. Przetwarzanie wstępne danych okulograficznych

W celu dokonania analizy uzyskanych danych użyto języka Python wraz z pakietami, takimi jak: Seaborn, Numpy, Pandas, Pyplot itp. Jak podają autorzy w licznych artykułach (metoda ta jest stosowana już od dawna) przy analizie danych okulograficznych warto pominąć pierwsze kilka sekund/minut eksperymentu (w zależności od czasu trwania całego badania). Jest to uzasadnione tym, że w czasie pierwszego skierowania wzroku na ekran monitora następuje akomodacja oka do zaistniałych warunków (np. do jasności ekranu monitora). W przypadku naszych eksperymentów usunięto pierwsze 3 s z pierwszego czarnego tła w eksperymencie u każdej z osób.

4.2. Uzyskane dane okulograficzne

4.2.1. Dane dotyczące całego eksperymentu

Po przeprowadzeniu eksperymentów z kalibracją oraz bez niej dokonano analizy uzyskanych danych okulograficznych. Najpierw przyjrano się mrugnieniom oraz fiksacjom wszystkich użytkowników (rys. 2).



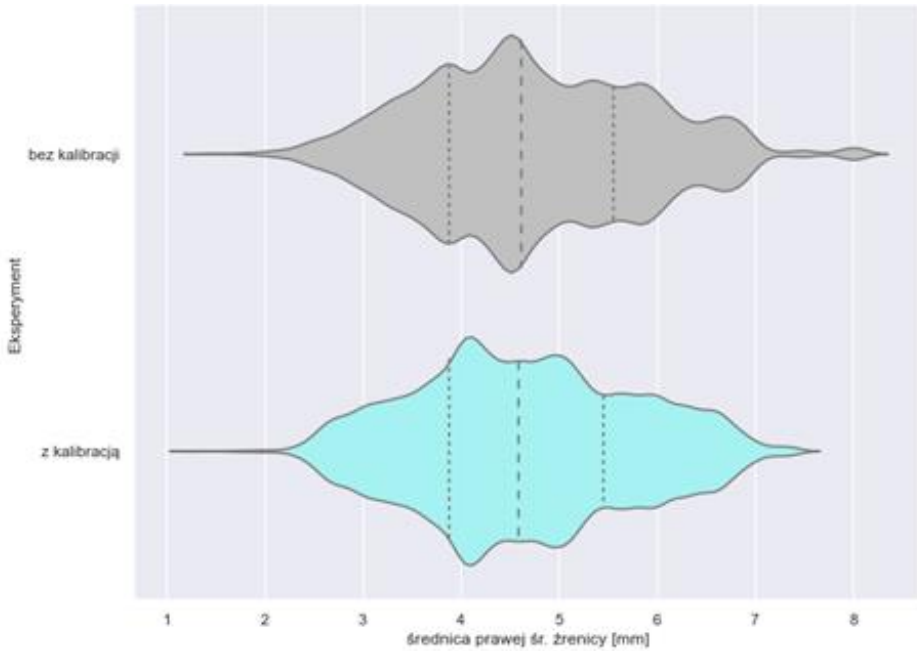
Rysunek 2. Wykresy pudełkowe dla wartości fiksacji i mrugnięć [s] [opracowanie własne]

Jak można zauważyć z powyższych wykresów pudełkowych, zarówno wartości fiksacji, jak i mrugnięć, niezależnie od przeprowadzenia kalibracji, przyjmowały niemal identyczny zakres. W dwóch typach eksperymentów mediana dla mrugnięć oraz mediana dla fiksacji przyjmowały niemal identyczne wartości, zgodne z literaturą. W obydwu typach eksperymentów zdarzały się wartości odstające (szary romb lub czarne zaznaczenie wielu rombów). Nie można jednak stwierdzić, że sam typ eksperymentu miał wpływ na uzyskane dane – raczej uzyskane wartości mrugnięć czy fiksacji mogły u niektórych osób przyjąć odrobinę większe wartości niż u reszty niezależnie od przeprowadzenia kalibracji.

W przypadku fiksacji, w eksperymencie z kalibracją, wartości ekstremalne przyjmowały nawet do 4 s. Prawdopodobnie sam typ eksperymentu nie miał tutaj znaczenia, ponieważ niektórzy po prostu dłużej skupiali się na danym punkcie na ekranie.

W poniższych rysunkach użyto wykresów skrzypcowych (ang. *violin plot*). Pozwalają one na wizualizację jednocześnie gęstości prawdopodobieństwa oraz rozkład dla kilku (tutaj dwóch) grup jednocześnie. Można powiedzieć, że dzięki takiemu typowi wykresu można odczytać na raz nie tylko informację na temat rozkładu danych (jak w przypadku wykresów pudełkowych), ale również informację odnośnie gęstości prawdopodobieństwa uzyskanych danych. Na poniższych wykresach zaznaczono pierwszy, drugi i trzeci percentyl za pomocą pionowych kresek na kolorowych obszarach.

Poniższe wykresy skrzypcowe pokazują wszystkie dane (dla całego eksperymentu trwającego w sumie 28 s).



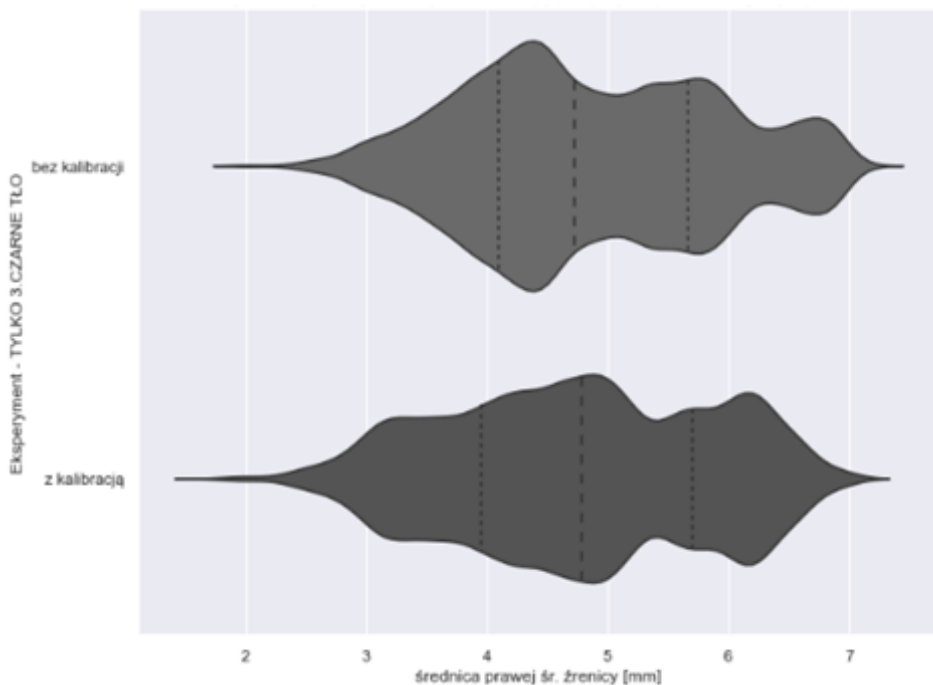
Rysunek 3. Wykresy skrzypcowe dla danych prawej średnicy źrenicy z całości danych dla dwóch typów eksperymentów [opracowanie własne]

Na powyższym zestawieniu widać, że zarówno pierwszy, jak i drugi percentyl, czyli mediana, są w obu przypadkach niemal identyczne. Jedynie trzeci percentyl jest minimalnie mniejszy w przypadku eksperymentu z kalibracją. Co jest istotne, to kształty uzyskanych wykresów skrzypcowych. Można w ten sposób porównać gęstość prawdopodobieństwa dla danych w obydwu typach eksperymentów. Okazuje się, że oba kształty są niemal identyczne. W przypadku eksperymentu bez kalibracji, można zaobserwować minimalnie różnice na końcach wykresu, jednak należy pamiętać, że dotyczy to wyłącznie pojedynczych próbek w całym zestawie danych.

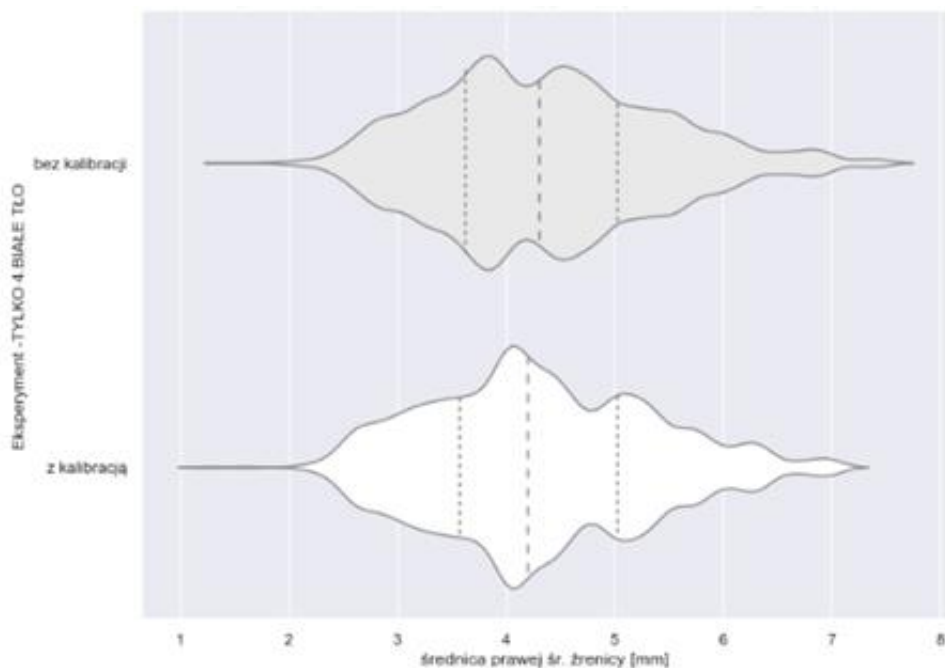
Wobec tego można uznać, że kalibracja nie wnosi istotnej jakości do uzyskanych danych. Należałoby to potwierdzić odpowiednim testem statystycznym.

4.2.2. Dane odnoszące się do poszczególnego typu tła

Została przeprowadzona analiza danych uwzględniających wyłącznie wielkość średnic prawych źrenic podczas każdego tła z osobna – czyli kolejno czarnego, białego, następnie znowu czarnego, białego i ostatniego czarnego (patrz rys. 1). Poniżej przedstawiono wykresy skrzypcowe dla wybranych dwóch typów tła.



Rysunek 4. Wykresy skrzypcowe dla danych prawej średnicy źrenicy z dwóch eksperymentów – pobrane tylko podczas oglądania czarnego tła (3. tło) [opracowanie własne]

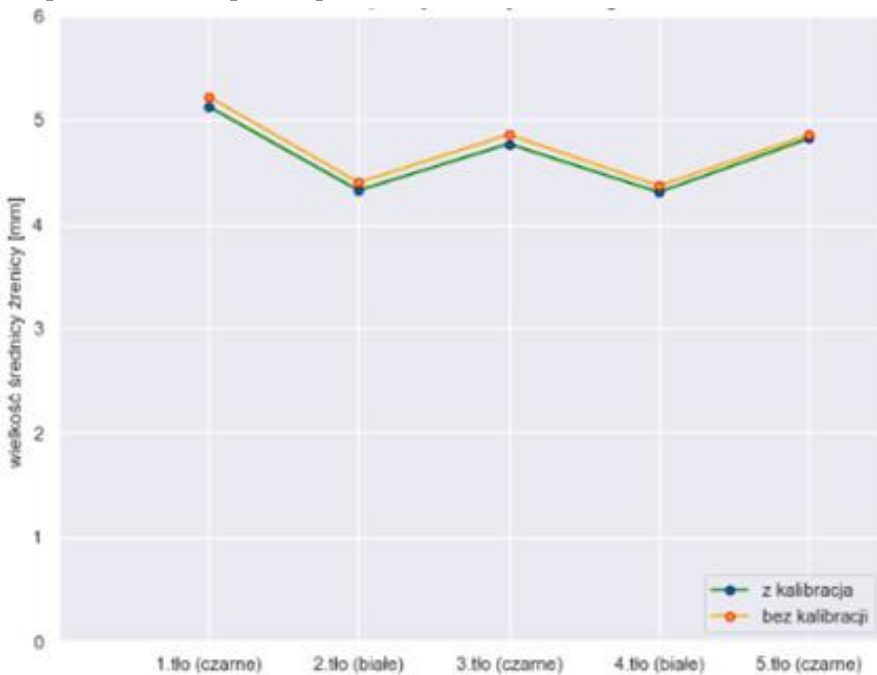


Rysunek 5. Wykresy skrzypcowe dla danych prawej średnicy źrenicy z dwóch eksperymentów – pobrane tylko podczas oglądania białego tła (4. tło) [opracowanie własne]

Należy zwrócić uwagę, że średnia wielkość średnicy źrenicy u użytkowników zmniejszyła się wraz z pojawieniem się białego tła: w przypadku tła czarnego wynosi ona ok. 4,8 mm, zaś w przypadku tła białego wynosi ona ok. 4,1 mm). Zmiana ta jest uzasadniona funkcją ochronną oka – w przypadku zmiany światła padającego do środka źrenicy na dużo jaśniejsze, następuje zwężenie się źrenicy, aby ochronić wnętrze oka.

Dla każdego zestawu wykresów można powiedzieć, że dane średnicy źrenicy przyjmują (niemal) identyczny zakres, jak i wartości mediany, pierwszego i trzeciego kwartylu (analizując każde tło z osobna). Nie można zauważyć żadnej znaczącej różnicy w uzyskanych danych, która wskazywałaby na to, który eksperyment został wykonany z kalibracją, a który bez.

Ostatnią z prezentowanych analiz jest wykres przedstawiający średnią wartość średnicy źrenicy dla poszczególnego wyświetlanego tła, z rozróżnieniem na typ eksperymentu. Na poniższym wykresie zobrazowano średnie obliczone ze wszystkich danych od użytkowników, wyszczególnione dla każdego tła osobno oraz każdego eksperymentu (z kalibracją oraz bez). Można zaobserwować znacznie widoczną różnicę w obliczonych średnich między tłem czarnym a białym (obojętnie którym z kolei), co świadczy o akomodacji oczu użytkowników do zaistniałego tła na monitorze. Natomiast dane pochodzące z eksperymentów z kalibracją przyjmują minimalne (pomijalnie) niższe wartości – minimalne różnice pomiędzy uzyskanymi danymi sięgają rzędu 0,1 mm. Należałoby jednak potwierdzić testami statystycznymi, że uzyskana różnica nie ma znaczenia statystycznego, czyli nie ma żadnego wpływu na jakość danych. Jest to istotna kwestia, ponieważ hipoteza o braku konieczności przeprowadzenia kalibracji staje się przez to bardzo prawdopodobna.



Rysunek 6. Wykresy średnich wartości prawej średnicy źrenicy dla każdego tła osobno, z rozróżnieniem na typ eksperymentu [opracowanie własne]

Jak widać z powyższych wykresów, mimo że próbka badawcza była stosunkowo mała (32 osoby), to uzyskane dane pokazują, że wyraźnie ujawnia się pewna tendencja, która potwierdza postawioną hipotezę. Z analizy uzyskanych danych w eksperymentach wynika, że dane pochodzące z eksperymentu, zarówno z kalibracją, jak i bez niej, nie różnią się. Oczywiście, żeby uwiarygodnić postawioną hipotezę, należałoby przeprowadzić badania na znacznie większej liczbie osób oraz na koniec potwierdzić testami statystycznymi. Jednak z przeprowadzonych badań można wywnioskować, że występuje duże prawdopodobieństwo, że kalibracja tego okulografu jest zbędna.

5. Wnioski

Niniejsze badania umożliwiły sprawdzenie, czy podczas korzystania z okulografu firmy Gazepoint GP3 w interfejsie komputer-człowiek można pominąć proces kalibracji. W tym celu wykonano dwa rodzaje eksperymentów, które przeprowadzono na 32 osobach – eksperyment z zastosowaną kalibracją przed pomiarem właściwym oraz eksperyment bez niej, z wykorzystaniem zmieniającego się kontrastowego tła. Z uzyskanych danych wynika wstępnie, że nie można zauważyć znaczącej różnicy pomiędzy danymi uzyskanymi z eksperymentów z kalibracją oraz bez niej. Z naszych obliczeń wynika, że w tych dwóch typach eksperymentów średnie oraz mediany z danych dotyczących średnicy źrenicy (analizując tła czarne i białe osobno) przyjmują niemal identyczne wartości. Również rozrzut danych fiksacji czy mrugnienie jest niemal identyczny, niezależnie od przeprowadzenia kalibracji. Uzyskane wykresy pudełkowe oraz skrzypcowe dotyczące średnicy źrenicy podczas poszczególnych typów tła sugerują, że kalibracja nie ma żadnego znaczącego wpływu na uzyskane dane. Oznacza to, że hipoteza o możliwości pominięcia kalibracji jest wysoce prawdopodobna, co należałoby potwierdzić jeszcze przeprowadzeniem badań z większą liczbą osób badanych oraz odpowiednimi testami statystycznymi.

Z wysokim prawdopodobieństwem można natomiast stwierdzić, że badania okulograficzne przeprowadzane nawet bez kalibracji mogą służyć jako wartościowy element w systemach wykorzystujących interakcję człowiek-komputer, na przykład do wspomagania działania agentów dialogowych.

6. Informacje o finansowaniu

Badania te były finansowane przez Centrum Badawcze Priorytetowego Obszaru Badawczego Sztuczna Inteligencja i Robotyka ze środków Politechniki Warszawskiej w ramach Programu Inicjatywa Doskonałości – Uczelnia Badawcza (IDUB).

Literatura

1. Stolecka-Makowska A., Wolny R., *Możliwości zastosowania techniki okulograficznej w ilościowych badaniach marketingowych*, Zesz. Nauk. Uniw. Ekon. w Katowicach, 195, 2014, s. 195-205.
2. Sielicka-Różyńska M., *Wykorzystanie badań eye trackingowych do oceny postrzegania wzrokowego znakowania produktów spożywczych*, Przem. SPOŻYWCZY, 12021, s. 18-23.
3. Bee N., André E., Tober S., *Breaking the ice in human-agent communication: Eye-gaze based initiation of contact with an embodied conversational agent*, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5773 LNAI, 2009.
4. Aracena C., Basterrech S., Snasel V., Velasquez J., *Neural Networks for Emotion Recognition Based on Eye Tracking Data*, in Proceedings – 2015 IEEE International

- Conference on Systems, Man, and Cybernetics, SMC 2015, s. 2632-2637 (Institute of Electrical and Electronics Engineers Inc., 2016), doi: 10.1109/SMC.2015.460.
5. Lim J.Z., Mountstephens J., Teo J., *Emotion Recognition Using Eye-Tracking: Taxonomy, Review and Current Challenges*, Sensors 20, 2020, s. 2384.
 6. Krejtz K., Duchowski A.T., Niedzielska A., Biele C., Krejtz I., *Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze*, PLoS One, 13, 2018.
 7. Chang K.-M., Wu Chueh M.-T., *Using Eye Tracking to Assess Gaze Concentration in Meditation*, Sensors 19, 2019, s. 1612.
 8. Sibanda W., Zhang Z., *Application of eye-tracking technology to predict concentration on HIV campaigns among students in South Africa*, in Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018, s. 988-994 (Institute of Electrical and Electronics Engineers Inc., 2018), doi: 10.1109/ASONAM.2018.8508415.
 9. Shintani T., Ishi C.T., Ishiguro H., *Analysis of role-based gaze behaviors and gaze aversions, and implementation of robot's gaze control for multi-party dialogue*, in HAI 2021 – Proceedings of the 9th International User Modeling, Adaptation and Personalization Human-Agent Interaction, s. 332-336 (Association for Computing Machinery, Inc, 2021). doi: 10.1145/3472307.3484653.
 10. Zygadło A., Kozłowski M., Janicki A., *Text-based emotion recognition in english and Polish for therapeutic chatbot*, Appl. Sci., 11, 2021.
 11. Chang K.M., Chueh M.T.W., *Using eye tracking to assess gaze concentration in meditation*, Sensors (Switzerland), 19, 2019.
 12. Peysakhovich V., Causse M., Scannella S., Dehais, F., *Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort*, Int. J. Psychophysiol., 97, 2015, s. 30-37.
 13. Oliva M., Anikin A., *Pupil dilation reflects the time course of emotion recognition in human vocalizations*, Sci. Rep., 8, 2018, s. 1-10.
 14. Bradley M.M., Miccoli L., Escrig M.A., Lang P.J., *The pupil as a measure of emotional arousal and autonomic activation*, Psychophysiology, 45, 2008.
 15. Berg A., *Wpływ alkoholu i środków odurzających na bezpieczeństwo uczestników ruchu drogowego*, in Kortowski Przegląd Prawniczy, 2017, wpia.uwm.edu.pl, <https://wpia.uwm.edu.pl/czasopisma/sites/default/files/uploads/KPP/2017/2/65-70.pdf> [data dostępu: 29.11.2022].

Badanie zasadności używania kalibracji w pomiarach okulograficznych w celu zastosowania w interfejsie człowiek-komputer

Streszczenie

Obecnie okulografia (ang. *eyetracking*) jest szeroko wykorzystywana zarówno w badaniach naukowych, jak i komercyjnych obejmujących wiele zróżnicowanych dziedzin, np. medycynę, edukację, psychologię, czy marketing. W niniejszym artykule zwracamy uwagę na jej zastosowanie w zakresie interakcji człowiek-komputer. Dzięki zastosowaniu okulografii można uzyskać informacje o stanie emocjonalnym (emocje pozytywne, negatywne) czy psychologicznym człowieka (np. koncentracja). Ważną rolę odgrywa tutaj bezinwazyjność stosowania okulografu, aby mógł on służyć jako dodatkowe źródło danych pochodzących od użytkownika podczas wykonywania badania. Kluczowym czynnikiem, niezależnie od dziedziny, jest brak ingerencji w komfort podczas pomiaru. Niestety obecnie nadal zalecane jest przez producentów okulografów przeprowadzenie kalibracji urządzenia przed każdym badaniem, co bywa kłopotliwe i mało komfortowe dla użytkownika. Celem przeprowadzonych badań było sprawdzenie, czy podczas korzystania z okulografu można pominąć proces kalibracji, aby pomiar okulograficzny stał się niezauważalny i nie miał negatywnego wpływu na samopoczucie użytkownika. W tym celu wykonano dwa rodzaje eksperymentów, które przeprowadzono na 32 osobach – eksperyment z zastosowaną kalibracją przed pomiarem właściwym oraz eksperyment bez niej, z wykorzystaniem zmieniającego się kontrastowego tła. Porównano uzyskane dane dotyczące m.in. wielkości średnicy źrenicy. Dokonane analizy pokazują, że nie można zaobserwować

znaczących różnic pomiędzy uzyskanymi danymi z tych dwóch typów eksperymentów. Jest to bardzo obiecujący wniosek, który z pewnością może służyć wielu badaniom, w szczególności w obszarze interakcji człowiek-komputer.

Słowa kluczowe: okulografia, kalibracja, interfejs człowiek-komputer

Investigating the relevance of calibration in eyetracking measurements for application in human-computer interface

Abstract

Nowadays, eyetracking is widely used in both academic and commercial research covering many different fields such as medicine, education, psychology, and marketing. In this paper we highlight its application in the field of human-computer interaction. Thanks to eyetracking methodology, it is possible to obtain information about the emotional (positive and negative emotions) or psychological state of a person (e.g. concentration). The non-invasiveness while using the eyetracker plays an important role here, so it can serve as an additional source of data coming from the user while performing examinations. A key factor, regardless of the field, is the users' comfort during the measurements. Unfortunately, it is currently still recommended by eyetracking manufacturers to calibrate the device before each examination, which can be troublesome and uncomfortable for the user. The aim of this study was to find out whether the calibration process can be omitted, so that the eyetracking procedure becomes unnoticeable and does not have a negative impact on the users' comfort. For this purpose, two types of experiments were performed on 32 subjects – an experiment with calibration applied before the actual measurement and an experiment without performing calibration, using a changing contrast background. The data obtained were compared with respect to, among other things, pupil diameter size. The data analysis show that no significant differences can be observed between the data obtained from these two types of experiments. This is a very promising finding that can certainly serve for many researches, especially in the area of human-computer interaction.

Keywords: eyetracking, oculography, calibration, human-computer interface

Indeks Autorów

Belka R.	93
Benke G.	35
Bier A.	114
Bujnik T.	134
Chodyka M.	124, 134
Czajka J.	26
Gabor-Siatkowska K.	148
Goc K.	35
Golovko V.	124, 134
Grzegórski K.	124
Janicki A.	148
Lentner M.	76
Leszczyńska-Sejda K.	35
Lichograj P.	124, 134
Łasica W.	52
Malarz J.	35
Mądrała A.	76
Ochmański M.	35
Pianowska K.	35
Różański S.A.	7
Sroczyński Z.	114
Sztandera J.	35