

Rozdział 6

Wykorzystanie charakterystyki dominujących częstotliwości harmonicznych i wyników przekształceń transformaty falkowej jako suplementacji atrybutów do analiz opisujących wpływ dźwięku na człowieka

Marek Kopczewski*

Streszczenie. Wzrost zainteresowania wpływem bodźców akustycznych na funkcjonowanie człowieka, zarówno w kontekście fizjologicznym, jak i psychologicznym, znajduje odzwierciedlenie w rozwijających się metodach badawczych, które są wykorzystywane do analizy wpływu dźwięku na sen oraz proces wybudzania się. Wskazując na rosnące możliwości sztucznej inteligencji w poszukiwaniu powiązań i korelacji między cechami zmienności częstotliwościowej dźwięku i wpływem na człowieka, w obecnych badaniach należy uwzględnić nowe parametry, które można wykorzystać w analizie. Autor pracy proponuje wykorzystanie wyników analizy sygnałów, jako suplementu do obecnych już metod badania wpływu dźwięku na człowieka. W analizowanych badaniach jako podstawę użyto ankiety, w której respondenci określali parametry snu, następnie przeprowadzono klasyfikację wyników oraz dodano dodatkowe parametry na podstawie analizy dźwięku. Wyniki badań wykazały zmianę w wynikach klasyfikacji.

Słowa kluczowe: analiza dźwięku, częstotliwości harmoniczne, transformata falkowa, sztuczna inteligencja

Wprowadzenie

Początki transformaty Fouriera przypadają na początek XIX wieku, kiedy to Joseph Fourier w pracy *analytique de la chaleur* (1822) zaproponował, że każda funkcja okresowa może zostać przedstawiona jako suma funkcji trygonometrycznych o różnych częstotliwościach, amplitudach i fazach. Koncepcja ta została zastosowana w analizie przewodnictwa ciepła. Podejście Fouriera opierało się na założeniu, że funkcje o dowolnym kształcie (również z punktami nieciągłości) można reprezentować za po-

* Wydział Informatyki, Politechnika Białostocka, Wiejska 45A, 15-351 Białystok, marek.kopczewski@pb.edu.pl

DOI 10.24427/978-83-68673-08-1_6

mocą nieskończonych szeregów trygonometrycznych. Idea ta spotkała się początkowo z krytyką środowiska naukowego, ponieważ istniały wątpliwości co do możliwości zastosowania takiej reprezentacji do funkcji niegładkich. Mimo tych kontrowersji, przedstawione podejście rozpoczęło rozwój nowej dziedziny matematyki – analizy harmonicznej. W drugiej połowie XIX wieku rozpoczęto formalizację teorii zaproponowanej przez Fouriera. Zostały opracowane warunki określające zbieżność szeregów Fouriera (warunki Dirichleta), a także poszerzono rozumienie funkcji i całek. Rozwój teorii funkcji rzeczywistych oraz przestrzeni funkcyjnych umożliwił precyzyjniejsze opisanie operacji związanych z analizą Fouriera. Wprowadzono całki Lebesgue’a, które umożliwiły analizę funkcji nieposiadających klasycznych całek Riemanna. Rozwijana równolegle teoria przestrzeni Hilberta pozwoliła na ujęcie szeregów Fouriera jako rozwinięcia w bazach ortonormalnych przestrzeni funkcyjnych. Dzięki temu transformata Fouriera została opisana jako przekształcenie liniowe w tych przestrzeniach. Transformata Fouriera, mimo swojej użyteczności, nie pozwala na uzyskanie informacji o lokalizacji czasowej składowych częstotliwościowych w sygnale. W celu częściowego rozwiązania tego problemu opracowano krótko czasową transformatę Fouriera (STFT), polegającą na zastosowaniu transformaty Fouriera do kolejnych fragmentów sygnału okienkowanego. STFT umożliwia częściową lokalizację informacji częstotliwościowej w czasie, jednak wiąże się z kompromisem pomiędzy rozdzielczością czasową a częstotliwościową. Zasada nieoznaczoności powoduje, że dokładniejsza lokalizacja w jednej dziedzinie skutkuje pogorszeniem dokładności w drugiej. Ograniczenia te doprowadziły do opracowania metod umożliwiających adaptacyjną analizę czasowo-częstotliwościową – w szczególności rozwinięto koncepcję transformacji falkowej, która umożliwia zmiennoskalową analizę sygnałów. W latach 80. XX wieku transformata falkowa została ujęta w formalne ramy matematyczne i przekształcona w narzędzie do zastosowań obliczeniowych. Jean Morlet wprowadził pojęcie falki jako funkcji o lokalnym charakterze czasowym i częstotliwościowym, służącej do analizy sygnałów zmiennych w czasie. Yves Meyer rozwinął pierwsze teoretyczne podstawy dla ciągłych i ortonormalnych baz falkowych. Ingrid Daubechies skonstruowała klasy falk o zwartym nośniku, spełniające warunki ortonormalności i posiadające określoną liczbę zanikających momentów. Falki te znalazły zastosowanie w dyskretniej analizie falkowej, w tym w przetwarzaniu sygnałów i obrazów. Stéphane Mallat opracował algorytm dekompozycji wielorozdzielczej (MRA), który umożliwia efektywną implementację dyskretniej transformaty falkowej. Algorytm ten pozwala na analizę sygnałów na różnych poziomach rozdzielczości i jest szeroko wykorzystywany w kompresji danych oraz redukcji szumów.

Wprowadzenie do transformaty Fouriera

Transformata Fouriera stanowi rozszerzenie idei szeregów Fouriera na funkcje nie-okresowe. Zamiast sumy dyskretnych częstotliwości, stosowana jest całka oznaczająca przejście do ciągłego widma częstotliwościowego. Dla funkcji $f(t)$ całkownej w sensie Lebesgue'a, transformata Fouriera definiowana jest wzorem:

$$\mathcal{F}\{f(t)\} = \mathcal{F}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i \omega t} dt \quad (1)$$

Transformacja ta umożliwia analizę zawartości częstotliwościowej sygnału bez odniesienia do jego reprezentacji czasowej. Znajduje zastosowanie w analizie widmowej, teorii systemów liniowych, rozwiązaniach równań różniczkowych, przetwarzaniu sygnałów oraz obrazów Weisstein (n.d.); Zieliński (2006). Analiza harmoniczna sygnałów umożliwia identyfikację ich składowych częstotliwościowych. W wielu zastosowaniach istotne jest określenie, które składowe częstotliwościowe (tzw. harmoniczne) są dominujące, tzn. mają największy udział w amplitudzie lub energii sygnału. Sygnał okresowy $x(t)$ o okresie T może być przedstawiony w postaci szeregu Fouriera:

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{i n \omega_0 t}, \quad \omega_0 = \frac{2\pi}{T} \quad (2)$$

Gdzie c_n to współczynniki Fouriera, a ω_0 to podstawowa częstość kątowna. Częstotliwości $n\omega_0$ przy których wartości $|c_n|$ są największe, uznawane są za częstotliwości harmoniczne dominujące.

W przypadku sygnałów nieokresowych, wykorzystywana jest transformata Fouriera:

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i \omega t} dt \quad (3)$$

Harmoniczne dominujące identyfikuje się jako te wartości ω , dla których $|X(\omega)|$ lub $|X(\omega)|^2$ przyjmują wartości największe. W praktyce, przy analizie sygnałów cyfrowych, stosowana jest dyskretna transformata Fouriera (DFT):

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-i \frac{2\pi}{N} kn}, \quad k = 0, 1, \dots, N-1 \quad (4)$$

Gdzie:

- $x[n]$ – próbki sygnału w dziedzinie czasu,
- N – liczba próbek,
- $X[k]$ – odpowiadające im współczynniki w dziedzinie częstotliwości.

Widmo amplitudowe sygnału określa się jako:

$$A[k] = |X[k]| \quad (5)$$

Moc widma:

$$P[k] = |X[k]|^2 \quad (6)$$

Do wskazania potencjalnej częstotliwości harmonicznej dominującej można zastosować następujące metody:

- **Detekcja maksimów lokalnych:**

$$A[k] > A[k-1] \wedge A[k] > A[k+1] \Rightarrow f_k \quad (7)$$

- **Kryterium progu amplitudy:**

$$A[k] > \alpha \cdot \max(A[k]), \quad \text{gdzie } 0 < \alpha < 1 \quad (8)$$

Przykładowo, dla $\alpha = 0,5$ identyfikowane są składowe przekraczające 50% maksymalnej amplitudy.

- **Energia względna danej harmonicznej:**

$$E_k = \frac{P[k]}{\sum_{j=0}^{N-1} P[j]} \cdot 100\% \quad (9)$$

Składowa jest uznawana za dominującą, jeżeli E_k przekracza zadany próg (np. 20%).

- **Porównanie z tłem widmowym (analiza SNR):**

$$\text{SNR}[k] = \frac{P[k]}{P_{\text{szum}}[k]} \quad (10)$$

gdzie $P_{\text{szum}}[k]$ – oszacowane tło widmowe.

Wprowadzenie do transformaty falkowej

Transformata falkowa jest metodą analizy sygnałów, umożliwiającą jednocześnie przedstawienie informacji w dziedzinie czasu i częstotliwości. Opiera się na przekształceniu sygnału z wykorzystaniem funkcji bazowych zwanych *falkami*, które mają ograniczony czas trwania i lokalizację częstotliwościową.

W przeciwieństwie do transformaty Fouriera, która dostarcza globalnego widma częstotliwościowego, transformata falkowa umożliwia analizę lokalną, co pozwala na uzyskanie zróżnicowanej rozdzielczości – dobrej w niskich częstotliwościach oraz dokładnej w czasie przy wysokich częstotliwościach Mallat (2008); Vetterli i Kovacevic (1995).

Transformata falkowa opiera się na analizie sygnału przy użyciu funkcji bazowych nazywanych *falkami*, które są skalowanymi i przesuwanymi wersjami tzw. falki matki $\psi(t)$. Ciągła transformata falkowa (CWT, ang. *Continuous Wavelet Transform*) wyraża się wzorem:

$$W(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \cdot \psi\left(\frac{t-b}{a}\right) dt \quad (11)$$

gdzie:

- $x(t)$ – analizowany sygnał,
- a – współczynnik skali (odwrotnie proporcjonalny do częstotliwości),
- b – współczynnik przesunięcia czasowego,
- $\psi(t)$ – funkcja falkowa (falka matka),
- $W(a, b)$ – współczynnik korelacji między sygnałem a falką na danej skali i w danym czasie.

Poniżej przedstawiono tabelę poszczególnych rodzin falek ze względu na ich typ oraz zastosowanie:

Tabela 6.1: Porównanie wybranych rodzin falek

Rodzina	Typ	Nośnik	Ortogonalność	Zastosowania
Daubechies (dbN)	DWT	Zwarty	Tak	kompresja, ogólna analiza
Symlet (symN)	DWT	Zwarty	Tak	analiza bez zniekształceń
Coiflet (coifN)	DWT	Zwarty	Tak	detekcja zmian trendu
Biorthogonal	DWT	Zwarty	Biortogonalne	rekonstrukcja sygnału
Morlet	CWT	Niezwarty	Nie	EEG, analiza widmowa
Mexican Hat	CWT	Niezwarty	Nie	detekcja impulsów
Haar	DWT	Zwarty	Tak	szybka, prosta dekompozycja

Transformata falkowa, szczególnie w postaci dyskretnej (DWT), umożliwia dekompozycję sygnału audio na pasma częstotliwościowe w sposób hierarchiczny. Dzięki temu możliwe jest wydzielenie sygnałów aproksymacyjnych i szczegółowych, co matematycznie zapisuje się jako:

$$x[n] = A_J[n] + \sum_{j=1}^J D_j[n] \quad (12)$$

gdzie:

- $A_J[n]$ – przybliżenie (komponenty niskoczęstotliwościowe) na poziomie J ,
- $D_j[n]$ – współczynniki falkowe (szczegółowe) na poziomie j w pozycji n ,
- J – liczba poziomów dekompozycji.

Z poziomów dekompozycji wyodrębnia się cechy takie jak energia falkowa w poszczególnych pasmach:

$$E_j = \sum_n |D_j[n]|^2 \quad (13)$$

gdzie E_j oznacza energię na poziomie j , a $D_j[n]$ – współczynniki szczegółowe dla tego poziomu.

Poziom złożoności w dziedzinie czasu i częstotliwości, bazujący na rozkładzie współczynników falkowych, określany jest za pomocą *entropii falkowej*, definiowanej wzorem:

$$H_j = - \sum_n p_j[n] \log_2 p_j[n], \quad \text{gdzie} \quad p_j[n] = \frac{|D_j[n]|^2}{\sum_k |D_j[k]|^2} \quad (14)$$

gdzie:

- H_j – entropia Shannona dla poziomu j , oparta na znormalizowanych energiach,
- $p_j[n]$ – znormalizowana energia, czyli udział energii danego współczynnika w całkowitej energii poziomu j ,
- $|D_j[n]|^2$ – energia danego współczynnika falkowego.

6.1. Studium przypadku

Celem przytoczonych badań było znalezienie, zbadanie i opisanie charakterystyki sygnału dźwiękowego o największej skuteczności wybudzania ze snu. Do badań została wytypowana grupa kobiet i mężczyzn w wieku od 25 do 75 lat pracująca umysłowo. Badanie zostało przeprowadzone w grupie 17 osób. Wszyscy badani otrzymali ankietę do własnoręcznego wypełniania przez okres 21 dni. Każda uzupełniona kolumna wraz z datą została określona jako oddzielna próbka – ankiety nie były rozpatrywane pod kątem określonych osób. W ankiecie zawarte zostały dane dotyczące snu, m. in. godzina zasypiania, godzina wybudzenia, stwierdzenie, czy była drzemka w ciągu poprzedniego dnia, ilość drzemek przed pełnym wybudzeniem się itp. Oprócz tego każda z osób odpowiedziała na kilka krótkich pytań dotyczących płci, wieku oraz obszaru zamieszkania.

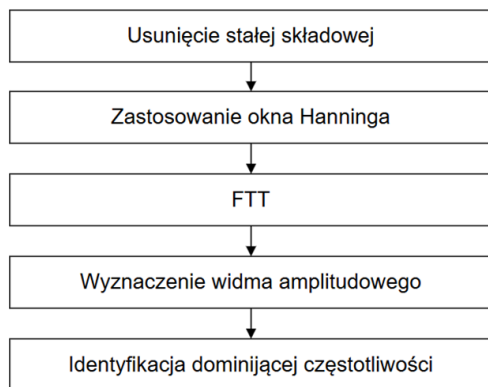
Klasyfikatorem dla badań był moment wybudzenia się. Czas do pełnego wybudzenia się był mierzony jako czas od zaśnięcia do momentu, w którym osoba ankietowana budziła się i zaczynała wykonywać poranne czynności.

Czas trwania badania wynosił 21 dni na osobę (21 próbek), co łącznie dało 351 próbek (w jednym przypadku uzupełniono 15 a nie 21 dni). Ankieta została przeprowadzona w okresie jesienno-zimowym, od początku października do końca grudnia.

Powyższe wyniki przekształceń zostały użyte jako suplementacja atrybutów danych do predykcji czasu snu oraz tzw. „przesypiania budzika” McFarlane, Garcia i Dyer (2021); McFarlane, Garcia, Verhagen i Dyer (2020) wykonanej przy użyciu następujących algorytmów: *Naive Bayes*, *Bayes Network*, *IbK*, *Regression* i *Random-Tree* Tan, Steinbach i Kumar (2003). Powyższe wyniki przekształceń zostały użyte jako suplementacja atrybutów danych do predykcji czasu snu oraz tzw. „przesypiania budzika” wykonanej przy użyciu następujących algorytmów: *Naive Bayes*, *Bayes Network*, *IbK*, *Regression* i *RandomTree*. Tan i in. (2003)

6.1.1. Ekstrakcja i wizualizacja dominujących częstotliwości harmonicznych na podstawie wybranych plików audio

Poniższy schemat prezentuje kroki wymagane w celu ekstrakcji dominującej częstotliwości harmonicznej w wymiarze czasu przy użyciu dyskretnej transformaty Fouriera.



Rysunek 6.1: Schemat blokowy identyfikacji dominującej częstotliwości

W pierwszym kroku z sygnału należy usunąć wartość średnią, co eliminuje składową stałą (DC). Dla sygnału $x[n]$ o długości N definiuje się wersję wyśrodkowaną jako:

$$x_c[n] = x[n] - \frac{1}{N} \sum_{k=0}^{N-1} x[k] \quad (14)$$

Aby ograniczyć przeciek widmowy związany z obcinaniem sygnału, stosuje się funkcję okna Hanninga. Jest to funkcja wagowa:

$$w[n] = 0,5 \left(1 - \cos \left(\frac{2\pi n}{N-1} \right) \right), \quad n = 0, 1, \dots, N-1 \quad (15)$$

Sygnal po zastosowaniu okna ma postać:

$$x_w[n] = x_c[n] \cdot w[n] \quad (16)$$

Następnie na sygnale $x_w[n]$ obliczana jest dyskretna transformata Fouriera (DFT), przechodząc w dziedzinę częstotliwości:

$$X[k] = \sum_{n=0}^{N-1} x_w[n] \cdot e^{-i \frac{2\pi kn}{N}}, \quad k = 0, 1, \dots, N-1 \quad (17)$$

Dla każdej częstotliwości obliczana jest wartość modułu amplitudy:

$$A[k] = |X[k]| = \sqrt{\text{Re}(X[k])^2 + \text{Im}(X[k])^2} \quad (18)$$

Analizie podlegają tylko wartości od $k = 1$ do $k = N/2$, ponieważ widmo jest symetryczne, a $k = 0$ odpowiada składowej stałej.

Ostatnim krokiem jest identyfikacja częstotliwości dominującej poprzez wyszukiwanie indeksu $k_{\max}$, dla którego $A[k]$ osiąga wartość maksymalną:

$$k_{\max} = \arg \max_{k \in [1, N/2]} A[k] \quad (19)$$

Na podstawie wyznaczonego indeksu $k_{\max}$ oblicza się częstotliwość dominującą w hercach według wzoru:

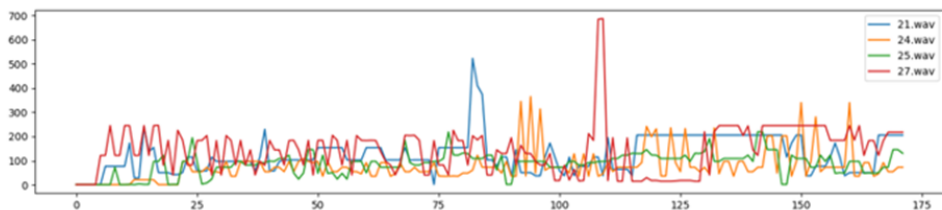
$$f_{\text{dom}} = \frac{k_{\max} \cdot f_s}{N} \quad (20)$$

Na poniższym diagramie przedstawiono wybrane dźwięki oraz ich częstotliwości harmoniczne dominujące (tj. najniższą częstotliwość dominującą wyrażoną w Hz) w funkcji czasu. Dla każdego fragmentu analizowanego sygnału czasowego wyznaczano dominującą częstotliwość z krokiem czasowym $\Delta t = \frac{1}{10}$ s.

Próbki oznaczone zostały numerami: 21, 24, 25 oraz 27.

Oś X na wykresie reprezentuje upływ czasu, natomiast oś Y – odpowiadającą częstotliwość dominującą. Przebieg przedstawia ewolucję najniższej składowej harmonicznego w czasie, co umożliwia identyfikację zmian w strukturze tonalnej analizowanych sygnałów.

Na Rysunku 6.3 zaprezentowano widma przedstawiające cztery spektrogramy sygnałów akustycznych, będące wizualizacją ich struktury czasowo-częstotliwościowej. Każdy z wykresów obrazuje ewolucję energii poszczególnych składowych częstotliwościowych w czasie, umożliwiając ocenę charakteru akustycznego sygnałów, ich harmoniczności, ciągłości oraz potencjalnego źródła dźwięku. Widma zostały przedstawione w kolejności zgodnej z Rysunkiem 6.2.



Rysunek 6.2: Schemat blokowy identyfikacji dominującej częstotliwości

Pierwszy spektrogram (próbka 21) charakteryzuje się silnym zagęszczeniem składowych częstotliwościowych, szczególnie w dolnym zakresie widma, oraz znaczną obecnością wyższych harmonicznych. Brak wyraźnych, stabilnych linii wskazuje na złożony charakter sygnału.

Drugi spektrogram (próbka 24) wykazuje znacznie wyraźniejszą strukturę harmoniczną. Widoczne są stabilne, poziome pasma (formanty), które przemieszczają się w czasie, co jest typowe dla dźwięków mowy ludzkiej lub śpiewu. Skoki w częstotliwości formantów mogą odpowiadać zmianom artykulacyjnym głosek lub melodiom intonacyjnym.

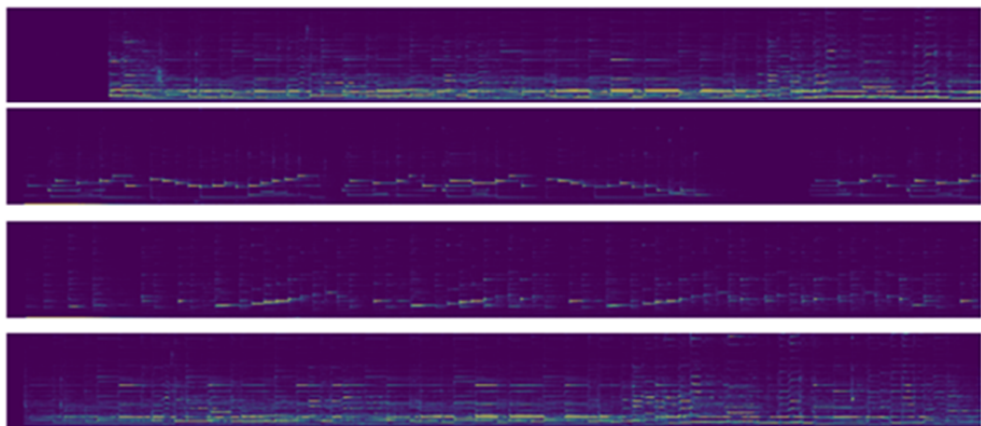
Trzeci spektrogram (próbka 25) prezentuje sygnał o podobnej strukturze harmoniczej, jednak bardziej płynnej i uporządkowanej niż w przypadku drugiego widma. Tego rodzaju przebieg może sugerować sygnał śpiewu lub instrumentu muzycznego emitującego dźwięki o regularnej strukturze.

Czwarty spektrogram (próbka 27) charakteryzuje się złożoną strukturą czasowo-częstotliwościową z obecnością licznych harmonicznnych rozciągających się zarówno w niskich, jak i wysokich częstotliwościach. Układ ten może świadczyć o nagraniu muzycznym zawierającym wiele współbrzmiących źródeł dźwięku – na przykład wieloinstrumentalnym utworze lub miksowanym materiale audio.

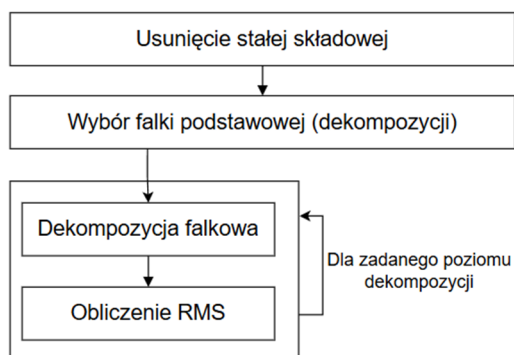
6.1.2. Ekstrakcja wyników i współczynników transformaty falkowej na podstawie wybranych plików audio

Proces przetwarzania sygnału obejmuje zastosowanie dyskretnej transformaty falkowej (DWT) do dekompozycji sygnału czasowego oraz obliczenie wartości skutecznej (RMS) na każdym poziomie aproksymacji. Przeprowadzono analizę z wykorzystaniem falki podstawowej HAAR (db1). Schemat działania przedstawiono na poniższym diagramie.

Dane wejściowe stanowił sygnał jednowymiarowy $x[n]$, próbkowany z częstotliwością $f_s = 10\text{Hz}$.



Rysunek 6.3: Widma przykładowych próbek



Rysunek 6.4: Schemat ekstrakcji współczynników dekompozycji falkowej

Funkcja falkowa Haar jest pierwszym i najprostszym przypadkiem ortonormalnej bazy falkowej. Dzięki minimalnemu wsparciu czasowemu oraz dwupunktowym filtrom rekonstrukcji i dekompozycji umożliwia wyjątkowo szybką dekompozycję sygnału, zapewniając przy tym wysoką rozdzielczość czasową. Jest szczególnie skuteczna w wykrywaniu transjentów – krótkotrwałych, niestacjonarnych zjawisk, takich jak impulsy perkusyjne, elementy artykulacyjne w mowie czy szybkie ataki instrumentów muzycznych.

Dekompozycję przeprowadzono rekurencyjnie do poziomu $L = 5$. W każdej iteracji uzyskiwano pary sygnałów aproksymacyjnych i detalicznych:

$$A_j[n] = \sum_k h[k] \cdot A_{j-1}[2n - k], \quad D_j[n] = \sum_k g[k] \cdot A_{j-1}[2n - k] \quad (20)$$

gdzie $A_0[n] = x_c[n]$, a $h[k]$ i $g[k]$ to odpowiednio filtr dolnoprzepustowy i górnoprzepustowy związany z wybraną funkcją falkową.

Na każdym poziomie aproksymacji A_j obliczono wartość skuteczną (RMS):

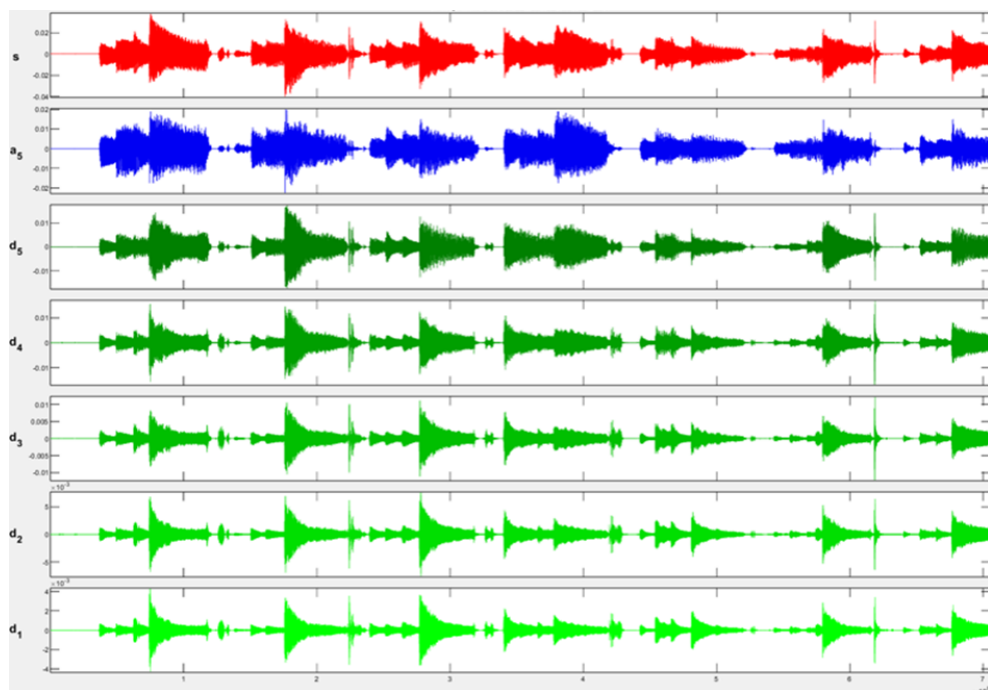
$$\text{RMS}(A_j) = \sqrt{\frac{1}{M_j} \sum_{n=0}^{M_j-1} A_j[n]^2} \quad (21)$$

gdzie M_j oznacza liczbę próbek sygnału aproksymacyjnego A_j na poziomie aproksymacji j .

Poniżej przedstawiono wyniki przekształceń w postaci diagramu, w którym na osi Y znajduje się amplituda sygnału, natomiast oś X reprezentuje upływ czasu.

Przebieg s odpowiada oryginalnemu sygnałowi wejściowemu $x[n]$. Przebiegi oznaczone jako d_1, d_2, d_3, d_4, d_5 przedstawiają szczegółowe komponenty (detale) sygnału wydobyte na poszczególnych poziomach dekompozycji. Przebieg a_5 odpowiada komponentowi aproksymacyjnemu uzyskanemu na najwyższym poziomie dekompozycji $L = 5$.

Diagram obrazuje hierarchiczną strukturę sygnału w dziedzinie czasu i częstotliwości, umożliwiając obserwację transjentów oraz głównych tendencji energetycznych w sygnale na różnych skalach.



Rysunek 6.5: Schemat ekstrakcji współczynników dekompozycji falkowej

Dla każdej jednostki czasu k zostały obliczone wartości aproksymacji i detali na przestrzeni całej próbki. Do ekstrakcji cech wykorzystano znormalizowane odchylenie standardowe, które posłużyło do kategoryzacji sygnałów.

Poziomy energii (RMS oraz aproksymacji) zostały zaprezentowane dla poziomu $L = 1$, natomiast poziomy energii detali przedstawiono w zakresie poziomów od $L = 1$ do $L = 5$.

Wyniki ekstrakcji współczynników dekompozycji falkowej, przedstawione w Tabeli 6.2, zostały pogrupowane na podstawie wartości odchylenia standardowego. Takie podejście umożliwiło wyodrębnienie grup sygnałów o zbliżonym poziomie zmienności energetycznej w dziedzinie czasu i skali. Grupowanie względem odchylenia standardowego pozwala na lepsze zrozumienie struktury sygnału oraz identyfikację obszarów o podwyższonej niestacjonarności.

Tabela 6.2: Grupowanie współczynników dekompozycji falkowej według wartości odchylenia standardowego

Próbka	RMS ($L = 1$)	ED1	ED2	ED3	ED4	ED5	EA ($L = 1$)
21	2	2	1	2	2	1	2
24	1	1	1	1	1	1	2
25	1	2	1	1	1	1	1
27	1	1	1	1	1	1	1

6.2. Badania

Celem przeprowadzonych badań była ocena przydatności analizy sygnału dźwiękowego jako dodatkowego wektora cech wspomagającego klasyfikację wyników ankiety dotyczącej zjawiska przesypania budzika. Tradycyjne podejścia oparte wyłącznie na danych deklaracyjnych, takich jak subiektywne odpowiedzi uczestników, mogą nie w pełni oddawać złożoność procesów percepcyjnych i fizjologicznych towarzyszących wybudzaniu.

W związku z tym w analizie uwzględniono również obiektywne cechy sygnału dźwiękowego, takie jak dominujące harmoniczne, poziomy RMS oraz energia falek. Dodatkowo przeprowadzono grupowanie sygnałów na podstawie współczynników uzyskanych z dekompozycji falkowej, co pozwoliło na identyfikację wzorców zmienności w dziedzinie czasu i skali. Takie rozszerzenie przestrzeni cech umożliwia bardziej precyzyjne modelowanie zależności między charakterystyką dźwięku a skutecznością wybudzania.

Tabela 6.3 prezentuje najlepsze wyniki klasyfikacji uzyskane przy zastosowaniu modelu *Bayes Network*. Na podstawie wyników ankiety można stwierdzić, że model sklasyfikował poprawnie 272 spośród 351 przypadków, co odpowiada dokładności na poziomie 77,49%. Liczba błędnych klasyfikacji wyniosła 79, co stanowi 22,51%

wszystkich obserwacji. Wartość współczynnika Kappa osiągnęła poziom 0,2821, co odzwierciedla zgodność klasyfikacji modelu z rzeczywistymi klasami z uwzględnieniem przypadkowej zgodności.

W zakresie błędów predykcji odnotowano następujące wartości:

- średni błąd bezwzględny (MAE): 0,334,
- pierwiastek średniego błędu kwadratowego (RMSE): 0,4095,
- średni błąd względny (MRE): 83,25%,
- względny błąd kwadratowy (RRSE): 91,12%.

Wszystkie te wskaźniki stanowią miary różnicy pomiędzy wartościami przewidywanymi przez model a rzeczywistymi obserwacjami, dostarczając informacji o precyzji predykcji modelu klasyfikacyjnego.

	Liczba próbek	Wynik
Liczba poprawnie sklasyfikowanych próbek	272	77.4929%
Liczba błędnie sklasyfikowanych próbek	79	22.5071%
Statystyka Kappa		0.2821
Błąd średni bezwzględny		0.334
Pierwiastek z błędu średniokwadratowego		0.4095
Względny błąd		83.9063%
Pierwiastek z względnego błędu średniokwadratowego		91.8502%
Łączna liczba próbek		351

Tabela 6.3: Wyniki klasyfikacji próbek na podstawie ankiety (Naive Bayes)

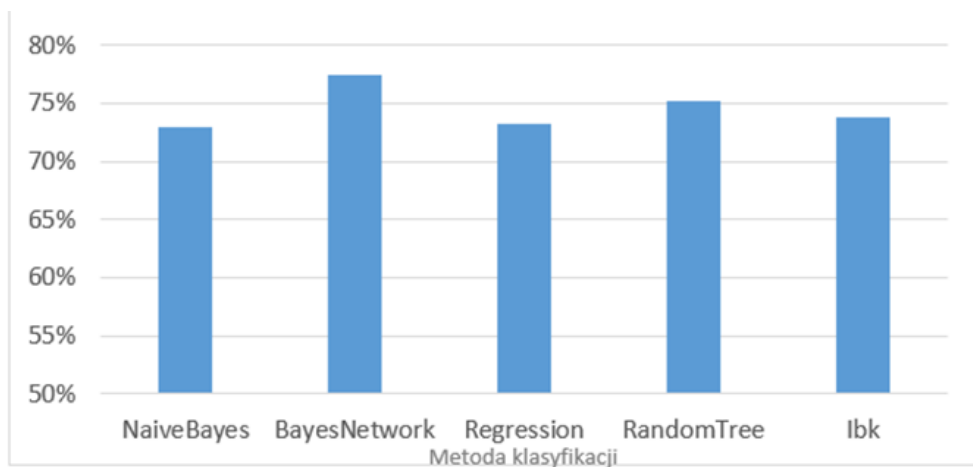
Wyniki wszystkich klasyfikatorów zaprezentowano na Rysunku 6.6, wszystkie metody klasyfikacji uzyskały zbliżone wyniki oscylujące w okolicach 75%.

Kolejnym krokiem badania było uwzględnienie wyników analizy częstotliwości dominujących oraz grupowanie dźwięków na trzy klasy:

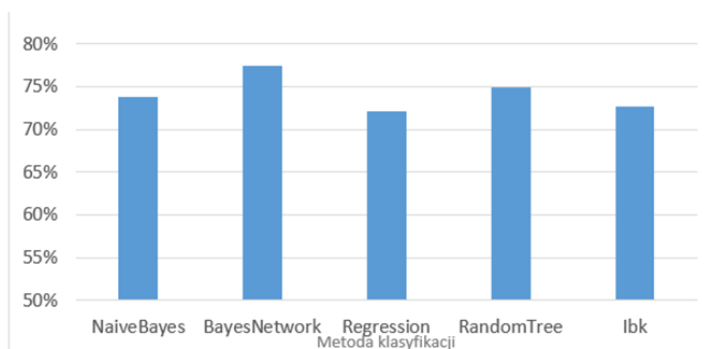
- **klasa 1** – sygnały zmienne okresowo,
- **klasa 2** – sygnały jednostajne,
- **klasa 3** – sygnały wysoko zmienne.

Następnie przeprowadzono ponowną klasyfikację próbek z uwzględnieniem przynależności sygnałów do jednej z powyższych klas. Wyniki poszczególnych klasyfikatorów przedstawiono na Rysunku 6.7.

Zastosowanie klasy sygnału jako dodatkowego wektora cech nie wpłynęło istotnie na ogólną skuteczność klasyfikacji. Najlepszy wynik uzyskany w poprzednim etapie pozostał bez zmian.



Rysunek 6.6: Wyniki poprawnej klasyfikacji dla poszczególnych klasyfikatorów na podstawie danych z ankiety



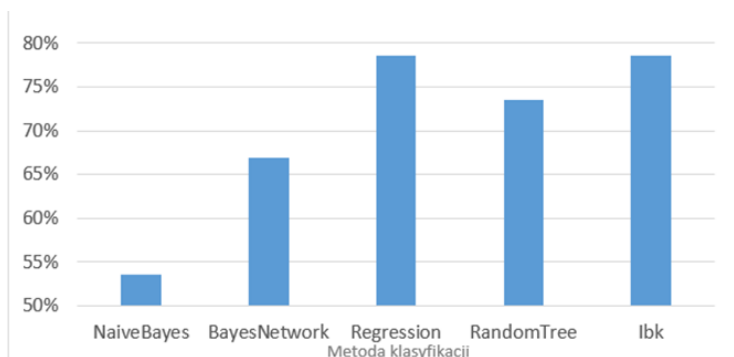
Rysunek 6.7: Wyniki poprawnej klasyfikacji dla poszczególnych klasyfikatorów na podstawie danych z ankiety oraz klasy sygnału

	Liczba próbek	Wynik
Liczba poprawnie sklasyfikowanych próbek	272	77.4929%
Liczba błędnie sklasyfikowanych próbek	79	22.5071%
Statystyka Kappa		0.2821
Błąd średni bezwzględny		0.334
Pierwiastek z błędu średniokwadratowego		0.4095
Względny błąd		83.9063%
Pierwiastek z względnego błędu średniokwadratowego		91.8502%
Łączna liczba próbek		351

Tabela 6.4: Wyniki klasyfikacji próbek na podstawie ankiety i klasy sygnału (Naive Bayes)

Jako kolejnego dodatkowe wektora cech użyto skategoryzowanych poziomów energii dla wyników transformaty falkowej, zostały uwzględnione (dla poziomów od $L=1$ do $L=5$): poziom RMS, poziom energii aproksymacji, poziom energii detali.

Na Rysunku 6.8 przedstawiono porównanie klasyfikatorów dla danych z dodatkowymi wektorami danych uzyskanych przy użyciu transformaty falkowej, zaobserwowano istotne różnice w wynikach względem poprzednich prób, zdecydowanie gorsze wyniki prezentują klasyfikatory Naive Bayes oraz Bayes Network, nieznacznemu pogorszeniu uległ klasyfikator RandomTree, natomiast wyniki klasyfikatorów Regression oraz lbk uległy nieznacznie poprawie.



Rysunek 6.8: Wyniki poprawnej klasyfikacji dla poszczególnych klasyfikatorów na podstawie danych z ankiety oraz klasy sygnału

Wyniki klasyfikacji zostały przedstawione w Tabeli. Wskazują na umiarkowaną skuteczność zastosowanego modelu. Spośród 351 przypadków, poprawnie sklasyfikowano 276, co przekłada się na dokładność na poziomie 78,63%. Współczynnik Kappa równy 0,4498 sugeruje umiarkowaną zgodność między przewidywaniami modelu a rzeczywistymi klasami, wyraźnie lepszą niż losowe przypisanie. Wartości błędów — średni błąd bezwzględny (0,2954) oraz pierwiastek średniego błędu kwadratowego (0,3846) — wskazują na umiarkowany poziom rozrzutu między wartościami przewidywanymi a rzeczywistymi.

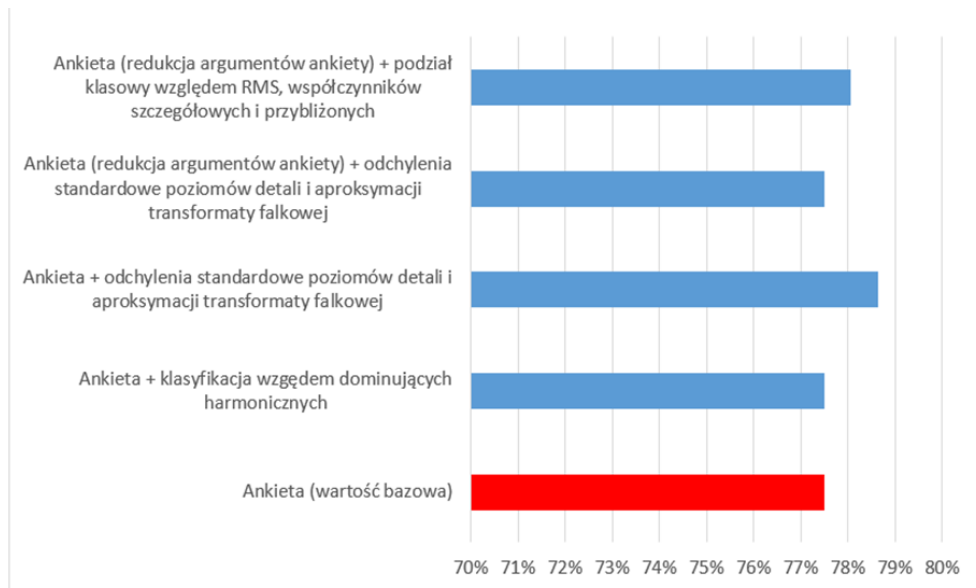
Na podstawie porównania zestawów wyników klasyfikacji można zauważyć wyraźną, choć umiarkowaną poprawę skuteczności modelu. W przypadku wyników uwzględniających wektor dodatkowy w postaci wyników transformaty falkowej liczba poprawnie sklasyfikowanych przypadków wzrosła z 272 do 276, co przełożyło się na wzrost dokładności z 77,49% do 78,63%. Jeszcze istotniejsza zmiana dotyczy współczynnika Kappa, który wzrósł z 0,2821 do 0,4498, wskazując na wyraźnie lepszą zgodność klasyfikacji z rzeczywistymi etykietami. Również wartości błędów uległy poprawie: średni błąd bezwzględny zmniejszył się z 0,334 do 0,2954, a pierwiastek średniego błędu kwadratowego z 0,4095 do 0,3846. Choć błędy względne

	Liczba próbek	Wynik
Liczba poprawnie sklasyfikowanych próbek	276	78.6325%
Liczba błędnie sklasyfikowanych próbek	75	21.3675%
Statystyka Kappa		0.4498
Błąd średni bezwzględny		0.2954
Pierwiastek z błędu średniokwadratowego		0.3846
Względny błąd		74.1949%
Pierwiastek z względnego błędu średniokwadratowego		86.2732%
Łączna liczba próbek		351

Tabela 6.5: Wyniki klasyfikacji próbek na podstawie ankiety i klasy sygnału (Regression)

nadal pozostają wysokie (spadek z 83,42% do 74,19% oraz z 91,16% do 86,27%), to obserwowane różnice sugerują, że wprowadzona modyfikacja modelu — przyniosła pozytywny efekt. Wnioski te wskazują na kierunek dalszych prac optymalizacyjnych, które mogą prowadzić do jeszcze lepszej jakości klasyfikacji.

Wyniki najlepszych klasyfikatorów dla wszystkich przeprowadzonych prób zaprezentowano na Rysunku 6.9. Na podstawie wyników można określić poprawę klasyfikacji próbek w przypadku użycia cech opartych o dekompozycję falkową.



Rysunek 6.9: Wyniki poprawnej klasyfikacji dla najlepszych klasyfikatorów na przestrzeni poszczególnych wektorów cech

Podsumowanie

W badaniu podjęto próbę wzbogacenia klasyfikacji odpowiedzi ankietowych dotyczących przesypiania budzika poprzez włączenie cech sygnału dźwiękowego. Tradycyjne metody, oparte wyłącznie na deklaracyjnych danych uczestników, często nie oddają pełnej złożoności procesów wybudzania. Dlatego autorzy zaproponowali podejście hybrydowe, łączące subiektywne dane z obiektywną analizą akustyczną, co pozwoliło na uchwycenie subtelnych różnic w percepcji i skuteczności działania alarmów.

Choć wprowadzenie dodatkowych cech nie przyniosło spektakularnego wzrostu skuteczności klasyfikacji, zauważono wyraźną poprawę spójności modelu i jakości predykcji. Wyniki te wpisują się w szerszy nurt badań nad multimodalnym podejściem do analizy zachowań użytkowników. Podobne strategie stosowano m.in. w badaniach nad detekcją senności kierowców, gdzie łączono dane EEG z analizą dźwięków otoczenia, czy w projektowaniu systemów alarmowych w środowiskach przemysłowych, gdzie uwzględniano parametry psychoakustyczne i fizjologiczne reakcje użytkowników. Wnioski z tych prac potwierdzają, że integracja danych sensorycznych z klasycznymi metodami deklaracyjnymi może prowadzić do bardziej wiarygodnych i użytecznych modeli predykcyjnych. Mahendra Prashanth i Sridar (2008); Skalski (2018)

W świetle tych obserwacji, badanie stanowi wartościowy krok w kierunku projektowania bardziej efektywnych systemów wybudzania, a także wskazuje potencjalne kierunki dalszych prac — zwłaszcza w zakresie personalizacji alarmów dźwiękowych i adaptacyjnych interfejsów użytkownika.

Bibliografia

- Mahendra Prashanth, K., i Sridar, V. (2008). The relationship between noise frequency components and physical, physiological and psychological effects of industrial workers. *Noise & Health*, 10(40), 90–98.
- Mallat, S. (2008). *A wavelet tour of signal processing: The sparse way*. Academic Press.
- McFarlane, S. J., Garcia, J. E. i Dyer, A. G. (2021). The awakening futures sound positive! commentary on the efficacy for audio to counteract sleep inertia. *Journal of Sleep Disorders & Therapy*, 10(1), 326.
- McFarlane, S. J., Garcia, J. E., Verhagen, D. S. i Dyer, A. G. (2020). Alarm tones, voice warnings, and musical treatments: A systematic review of auditory countermeasures for sleep inertia in abrupt and casual awakenings. *Clocks & Sleep*, 2(4), 416–433.
- Skalski, M. (2018). *Zaburzenia snu w codziennej praktyce* (2nd ed.). Medical Tribune Polska.

- Tan, P.-N., Steinbach, M. i Kumar, V. (2003). *Introduction to data mining*. University of Minnesota. (Lecture materials)
- Vetterli, M., i Kovacevic, J. (1995). *Wavelets and subband coding*. Prentice Hall.
- Weisstein, E. W. (n.d.). *Fast fourier transform*. MathWorld – A Wolfram Web Resource. (Retrieved from <https://mathworld.wolfram.com/FastFourierTransform.html>)
- Zieliński, T. P. (2006). *Cyfrowe przetwarzanie sygnałów. od teorii do zastosowań*. Wydawnictwo Komunikacji i Łączności.