

**Recenzja rozprawy doktorskiej pt.  
„Wykorzystanie metod nierównomiernego próbkowania i inteligentnych algorytmów  
przetwarzania sygnałów do rekonstrukcji archiwalnych nagrań dźwiękowych”  
mgr inż. Piotra Kardasza**

**1. Jaki jest problem naukowy (teza) rozprawy i czy został on trafnie i jasno sformułowany? Jaki charakter ma rozprawa?**

Celem rozprawy jest selekcja i optymalizacja metod i algorytmów przetwarzania sygnałów w związku z zagadnieniem rekonstrukcji archiwalnych nagrań. Teza rozprawy wskazuje algorytmy nierównomiernego próbkowania i wymienione już w tytule „inteligentne” algorytmy. Jest sformułowana dość ogólnikowo, co ma miejsce w większości rozpraw doktorskich. Przydałaby się jednak definicja algorytmu inteligentnego, pojęcie to nie jest ściśle określone. W pracy odnosi się ono do sieci neuronowych, algorytmów ewolucyjnych, genetycznych, logiki rozmytej. Nie budzi to sprzeciwu, jednak przydałoby się przedstawienie jakiegoś kryterium podziału algorytmów na inteligentne i „nieinteligentne”.

Zagadnienia, którymi zajął się autor rozprawy, są istotne z punktu widzenia zabezpieczenia i wykorzystania dźwiękowych zasobów archiwalnych. Problemem interesują się nie tylko archiwiści, ale również osoby prywatne, które zgromadziły kolekcję nagrań płytowych lub zapisanych na taśmie magnetycznej. Osoby te często nie posiadają niezbędnej wiedzy z dziedziny przetwarzania sygnałów, interesuje je oprogramowanie, które umożliwi detekcję zakłóceń i automatycznie dobierze algorytmy ich redukcji. Stworzeniem takiego narzędzia zajął się autor rozprawy. Ma ona charakter eksperymentalno – teoretyczny, z naciskiem na to pierwsze słowo.

**2. Czy w rozprawie przeprowadzono właściwą analizę źródeł? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?**

Bibliografia liczy 97 pozycji, 10 z nich to publikacje własne autora rozprawy, zamieszczone w periodykach krajowych. Moim zdaniem, bibliografia zawiera podstawowe pozycje odnoszące się do tematu rozprawy. W rozprawie nie ma osobnej części poświęconej wyłącznie analizie źródeł. Analiza ta jest rozproszona, jej elementy znaleźć można w kolejnych rozdziałach. Część opisowa rozprawy jest rozbudowana, jej lektura nie budzi wątpliwości, że autor dobrze opanował teorię przetwarzania sygnałów fonicznych, w szczególności powstawanie zakłóceń i metody ich tłumienia. Mgr inż. Piotr Kardasz posiada również rozległą wiedzę praktyczną – zna właściwości nośników zapisu analogowego (p.2.3) i związane z nimi zakłócenia (rozdział 3). Większość pozycji bibliograficznych pojawiła się w ostatnim dziesięcioleciu, co świadczy o dobrej orientacji w stanie wiedzy światowej w dziedzinie przetwarzania archiwalnych zasobów dźwiękowych<sup>1</sup>. Na uwagę zasługuje przedstawienie podstaw tzw. „rzadkiej” aproksymacji, w tym oszczędnego próbkowania (p.2.4).

---

<sup>1</sup> Autor cytuje książkę prof. T.P.Zielińskiego z 2005r, jednak bardziej adekwatne byłoby odwołanie się do książki „Cyfrowe przetwarzanie sygnałów w telekomunikacji, podstawy, multimedia, transmisja” z 2014r.

### 3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej metody i czy założenia są uzasadnione?

Celem autora było stworzenie oprogramowania, umożliwiającego klasyfikację archiwalnych nagrań pod kątem występujących zakłóceń oraz automatyczny dobór algorytmów służących do ich stłumienia i poprawy jakości nagrania. Moim zdaniem, cel ten został osiągnięty.

Proces rekonstrukcji rozpoczyna się od uruchomienia klasyfikatora nagrań. Na Rys.4.5 przedstawiono schemat klasyfikatora. Zawiera on algorytmy filtracji pasmowej, oceny odstępów sygnału od szumu w podpasmach, obliczenia pasma sygnału, identyfikacji najczęściej występujących zakłóceń harmoniczných (np. przydźwięk sieci, pilot analogowego sygnału stereofonicznego), detekcji zakłóceń impulsowych. Proces klasyfikacji ma kluczowe znaczenie, gdyż nieprawidłowy wynik klasyfikacji może spowodować uruchomienie niewłaściwych (dla danego nagrania) algorytmów rekonstrukcji. Szkoda jedynie, że ten algorytm został opisany w skrócie, brak jest dokładnych danych dotyczących zastosowanej sieci neuronowej, jej uczenia, przebiegu i wyników testowania. W testach wykorzystano podobno 100h nagrań, tym bardziej warto byłoby przedstawić wyniki testowania, w szczególności stopę błędów klasyfikacji itp.

Kolejnym etapem rekonstrukcji jest ujednoczenie częstotliwości próbkowania. Wybrano częstotliwość 192 kHz i opracowano efektywny algorytm zmiany częstotliwości próbkowania z 44.1 kHz na częstotliwość docelową. Opisany w p.4.2 algorytm polega na zastosowaniu stosunkowo niskiej częstotliwości pośredniej, interpolacji liniowej i filtracji końcowej. Jest to właściwe i pomysłowe rozwiązanie, umożliwiające uzyskanie dobrej jakości dźwięku przy małej złożoności obliczeniowej algorytmu.

Wiele uwagi poświęca autor rozprawie zagadnieniu identyfikacji składowych sinusoidalnych sygnału akustycznego (p.4.4). Aby uniknąć wyznaczania transformaty DFT w rozległym oknie, a co za tym idzie, obliczania wielkiej liczby współczynników transformaty niemających związku z zakłóceniem, proponuje on zastosowanie metod „rzadkiej” aproksymacji. W szczególności jest to algorytm ewolucyjny (p.4.4.2) i algorytm selektywnego obliczania współczynników STFT (p.4.4.3). W tym kontekście autor odwołuje się do zagadnienia „oszczędnego” próbkowania, czyli estymacji parametrów sygnału na podstawie ograniczonej liczby losowo pobranych próbek. Opracowany przez autora rozprawie ewolucyjny algorytm detekcji składowych sinusoidalnych wykorzystuje to „oszczędne” próbkowanie. Testy wykazały jego dużą skuteczność w przypadku, gdy mamy do czynienia z pojedynczym tonem zakłóconym szumem. Dobre wyniki estymacji amplitudy, częstotliwości i fazy sygnału sinusoidalnego zakłóconego szumem (SNR=20 dB) otrzymano na podstawie 128 wylosowanych próbek (na 2048 wygenerowanych). Szkoda, że nie przetestowano algorytmu dla kombinacji liniowej kilku sygnałów harmoniczných, a przede wszystkim, dla nagrań archiwalnych. Z drugiej strony, w nagraniach pojawiają się wprawdzie utracone segmenty, np. w miejscach wystąpienia zakłóceń impulsowych, jednak zdecydowana większość próbek jest dostępna, a więc estymacja parametrów sinusoidy na podstawie bardzo niewielkiej liczby próbek nie jest tu konieczna.

W p.4.4.3 autor przedstawia algorytm selektywnej analizy widmowej, opartej na obliczaniu współczynników STFT dla zestawu z góry określonych częstotliwości. Podejście takie ma duże zalety w przypadku, gdy znane są częstotliwości zakłóceń, np. przydźwięk sieci i jego harmoniczne. Pozwala uniknąć obliczania DFT dla segmentów sygnału liczących setki tysięcy próbek. Ponadto znakomicie nadaje się do implementacji na procesorach pracujących równolegle (p.4.4.4). Estymator amplitudy składowej sinusoidalnej sygnału podano we wzorze (4.23) – ma on związek z częścią urojoną współczynnika STFT. Analogicznie, zastępując funkcję sin funkcją cos, otrzymuje się estymator amplitudy składowej

cosinusoidalnej – czyli części rzeczywistej współczynnika STFT. Ciekawe byłoby zbadanie właściwości tych estymatorów – szczególnie ich efektywności. Czy są to estymatory optymalne w sensie Cramera – Rao? Analiza odwołująca się do pracy D. C. Rife'a i R. R. Boorstyna podniosłaby wartość rozprawy <sup>2</sup>.

Opracowany algorytm selektywnego obliczania współczynników STFT wykorzystano w algorytmie redukcji harmonicznym zakłóceń quasistacjonarnych (p.4.5) i w algorytmie kompensacji zmian prędkości przesuwu nośnika (p.4.6). Istotą tego pierwszego algorytmu jest zastąpienie filtracji odejmowaniem estymaty zakłócenia, co jest jego oryginalną cechą.

Proponowana przez mgr inż. Piotra Kardasza redukcja szumów opiera się na dekompozycji sygnału na część harmoniczną i aperiodyczną (p.4.7). Sama redukcja szumu polega na zastosowaniu tłumienia szumu w 10 podpasmach częstotliwości. Prezentacja widma sygnału (Rys.4.22) świadczy o prawidłowym działaniu algorytmu, jednak, moim zdaniem, brakuje tu wyników badań odsłuchowych.

Podobny charakter ma algorytm redukcji zakłóceń impulsowych – tu również nastąpiło oddzielenie części harmonicznym od aperiodycznej i zastosowany został prosty algorytm progowy wykrywania i eliminacji impulsów z części aperiodycznej.

Autor rozprawy zaproponował kilka algorytmów służących do poszerzania pasma sygnałów, które zostały nagrane na nośniku o słabej jakości (p.4.9). Algorytmy te bazują na rekonstrukcji wyższych harmonicznym poprzez ich syntezę lub przekształcenie nieliniowe. Autor dowodzi eksperymentalnie, że widmo istotnie ulega poszerzeniu, wyraża też pogląd, że jakość sygnału ulega poprawie. Moim zdaniem, brakuje tu badań odsłuchowych, nawet nieformalnych. Pozytywny wynik takich badań byłby istotnym argumentem za przyjęciem do realizacji sprzętowej proponowanych algorytmów poszerzania pasma sygnału.

W rozdz.5 rozprawy autor przedstawia realizację praktyczną wymienionych algorytmów w formie gotowego do działania uniwersalnego programu. Rozważa również możliwości implementacji sprzętowej oprogramowania, m.in. w układach FPGA. Świadczy to o jego wysokich kompetencjach inżynierskich.

#### **4. Na czym polega oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy reprezentowanego przez literaturę światową?**

Mgr inż. Piotr Kardasz opublikował 10 artykułów w czasopiśmie krajowych, są to prace samodzielne. Do oryginalnych pomysłów autora rozprawy można, moim zdaniem, zaliczyć:

- Efektywny algorytm zmiany częstotliwości próbkowania, pozwalający na obniżenie częstotliwości pośredniej. Połączono tu dwa podejścia: wstawianie próbek zerowych na wysokiej częstotliwości pośredniej i interpolację wartości próbek.
- Zastosowanie algorytmów „rzadkiej” aproksymacji do identyfikacji składowych sinusoidalnych: algorytm ewolucyjny i selektywny pomiar współczynników STFT.
- Zastąpienie filtracji odejmowaniem estymat zakłóceń harmonicznym w algorytmie redukcji harmonicznym zakłóceń quasistacjonarnych.
- Wydzielenie składowej aperiodycznej sygnału fonicznego w celu redukcji szumu i zakłóceń impulsowych.
- Algorytm automatycznej klasyfikacji nagrań dla potrzeb ich rekonstrukcji.

---

<sup>2</sup> D. C. Rife and R. R. Boorstyn, „Single-tone parameter estimation from discrete-time observations”, IEEE Transactions on Information Theory, vol. 20, pp. 591-598, September 1974

## **5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia wyników? (redakcja rozprawy)**

Struktura rozprawy (układ rozdziałów i podrozdziałów) jest właściwa, podporządkowana celowi rozprawy. Język i styl są poprawne, występują tylko nieliczne błędy interpunkcyjne. Drobniejsze problemy natury redakcyjnej i merytorycznej przedstawiłem w Dodatku.

W tym punkcie należałoby wspomnieć niezbyt fortunne wprowadzenie podziału zniekształceń nieliniowych na analogowe i cyfrowe (str.7), gdyż w końcu odnoszą się one do sygnału analogowego, odbieranego przez słuchacza.

Autor kilkakrotnie odwołuje się do własnych publikacji, w których czytelnik ma odnaleźć szczegóły algorytmów (np. na str. 88 jest odwołanie do pracy [48], na str. 107 do pracy [46], na str. 110 do [51], na str. 120 do [42]). Cytowania są oczywiście potrzebne, jednak szczegóły algorytmów opracowanych przez autora winny się znaleźć również w tekście rozprawy.

Lekturę rozprawy ułatwiłby wykaz oznaczeń i skrótów.

## **6. Jakie są wady i słabe strony rozprawy?**

Jak już wspomniałem, za mało jest w tej pracy wyników badań odsłuchowych. Odnosi się to szczególnie do algorytmów regeneracji brakujących harmonicznych w celu poszerzenia pasma sygnału. Autor wykazuje, że pasmo ulega poszerzeniu, natomiast nie wiadomo, czy efekt ten jest pozytywnie oceniany przez słuchaczy. Podobne zastrzeżenie można mieć do prezentacji wyników działania algorytmów redukcji zakłóceń impulsowych i szumu. Moim zdaniem, przeprowadzenie, choćby nieformalnych badań odsłuchowych byłoby bardzo pożądane, szczególnie odnosi się to do algorytmu poszerzania pasma.

Autor stwierdza, że do testowania systemu rekonstrukcji nagrań wykorzystał 100h materiału dźwiękowego, jednak w rozprawie nie zamieszczono dokumentacji tych badań. Należało to uczynić, nawet w dodatku do rozprawy.

## **7. Jaka jest przydatność rozprawy dla nauk technicznych?**

Mgr inż. Piotr Kardasz opracował i zrealizował system automatycznej rekonstrukcji archiwalnych nagrań. Użytkownik systemu nie musi posiadać wiedzy z dziedziny przetwarzania sygnałów, gdyż kolejne kroki: klasyfikacja nagrania, detekcja i korekcja zakłóceń wykonywane są automatycznie. Oprogramowanie to mogłoby ułatwić pracę archiwistów, ale przede wszystkim osób prywatnych, posiadających zbiory płyt lub taśm w domowych archiwach.

Algorytmy „rzadkiej” aproksymacji, zastosowane do analizy widmowej, mogłyby być przedstawione w publikacji o szerszym zasięgu, zainteresowałyby specjalistów z dziedziny przetwarzania sygnałów.

## **8. Do której kategorii recenzent zalicza rozprawę?**

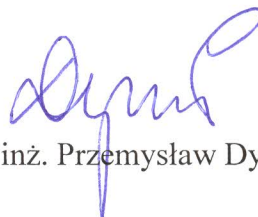
Rozprawę zaliczam do kategorii: spełniająca wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy.

## Dodatek: Wykaz zauważonych usterek redakcyjnych i nieścisłości (w tym sprawy do wyjaśnienia)

Wymienione tu drobne usterki w redakcji rozprawy nie mają wpływu na jej wartość merytoryczną. Do wyjaśnienia pozostaje problem identyfikacji harmonicznych (p.4.5.2).

- Błędy interpunkcyjne: str. 18, 57, 142, 144
- Na str. 18/19 dwukrotnie pojawia się ten sam akapit.
- Wyniki działania programu w formie listingu (Rys.2.9) nie są zbyt przejrzyste, wystarczyłby komentarz.
- We wzorze na odwrotną transformatę Fouriera (2.2) jest błąd (znak minus). Ten sam błąd pojawia się we wzorze (2.4) na odwrotną dyskretną transformatę Fouriera. Ponadto, we wzorach (2.1) i (2.2) stosuje się podwójne oznaczenie transformaty Fouriera:  $X(j\omega)$  i  $X(\omega)$ .
- W Tab. 2.2 jest mowa o oknie Hanna i Hanninga. W obu przypadkach chodzi o Juliusa von Hanna, więc raczej należałoby pisać o dwóch (bardzo do siebie zbliżonych) odmianach tego okna.
- We wzorze (2.26) wektor  $f$  występuje jako wiersz. Czy z uwagi na (2.27) nie powinien być kolumną? Z kolei na str.44 stwierdza się, że  $G$  jest macierzą kwadratową. Moim zdaniem, chodzi tu o  $\Phi$ .
- Czy wzory (4.34) i (4.35) służą do obliczenia amplitud  $A_{\sin}$  i  $A_{\cos}$ ? Jeśli tak, to brakuje znaków równości. Ponadto, nie zdefiniowano ciągów  $f_0(n)$  i  $S(nT)$ .
- Na str.103 autor twierdzi, że całka  $\int_0^{2\pi} \sin^2(t) dt$  nie ma wartości, podczas gdy tę wartość posiada i wynosi ona  $\int_0^{2\pi} \sin^2(t) dt = \pi$ . Podobnie całka (4.29) istnieje dla  $a=1$  i dowolnego  $b \neq 0$ . Nawiasem mówiąc, nie jest dla mnie jasne, czemu służy publikacja wzoru (4.27) i Rys.4.13 odnoszących się do wartości całki  $\int_0^{2\pi} \sin(t) \sin(at) dt$ . Okres funkcji podcałkowej zależy od wartości parametru  $a$ , w ogólnym przypadku funkcja podcałkowa  $\sin(t)\sin(at)$  może nie być funkcją okresową. Wartość całki zależy od przedziału całkowania, dlatego autor wybrał zakres od 0 do  $2\pi$ ? Całkując od 0 do np.  $20\pi$  można łatwo się przekonać, że maksimum wartości całki już nie występuje dla wartości parametru  $a=0.96$  (jak to sugeruje autor), tylko dla  $a \approx 1$ .
- Wzór (4.38) wymagałby uzupełnienia wskazującego, że jest on stosowany jedynie wówczas, gdy moc sygnału przekracza moc szumu, w przeciwnym wypadku  $k=0$ . Jest komentarz pod wzorem, jednak wypadałoby to ograniczenie zapisać w samym wzorze.

- Na str. 119 autor odwołuje się do podrozdziału 4.6.2, który nie istnieje. Chodzi zapewne o 4.7.2.



dr hab. inż. Przemysław Dymarski