

dr inż. Marcin Michałek

ekspert w Pracowni Analizy Mowy i Nagrań Instytutu Ekspertyz Sądowych

Wybrane metody redukcji zakłóceń utrwalonych w dowodowych nagraniach dźwiękowych

Wstęp

Jednym z podstawowych badań fonoskopijnych jest analiza audytywna nagrania wraz ze spisaniem treści utrwalonych w nim wypowiedzi [1–5]. Doświadczenie eksperckie w tej dziedzinie wskazuje, że w dźwiękowych nagraniach dowodowych bardzo często występują zakłócenia, przeważnie o zmiennej charakterystyce, znacznej amplitudzie i w paśmie częstotliwościowym pokrywającym się z pasmem sygnału mowy. Zakłócenia mogą wynikać ze sposobu nagrywania, warunków akustycznych podczas zdarzenia, jak również z użytej techniki rejestracji, a także jakości urządzeń i nośników [1, 6, 7].

Niska jakość zapisu powoduje konieczność stosowania korekcji nagrania w celu zwiększenia zrozumiałości i wyrazistości badanych sygnałów, w tym także sygnału mowy [1, 3, 4]. Korekcja taka odbywa się z wykorzystaniem narzędzi do zmiany parametrów nagrania, filtracji sygnałów oraz redukcji zakłóceń i zniekształceń [8–10]. Z uwagi na nierzadko odległe terminy wykonania ekspertyz fonoskopijnych i obszerność materiału dowodowego przekazywanego do badań organ prowadzący postępowanie, zwłaszcza na jego wczesnym etapie, może zlecić wykonanie kopii dowodowych zapisów wraz z korekcją jakości. Umożliwia to zarówno zapoznanie się z ich treścią, jak i ewentualne wskazanie, które konkretnie zapisy należy objąć dalszymi szczegółowymi badaniami, co może przyczynić się do skrócenia czasu wykonywania ekspertyzy. Ograniczenie zakłóceń w dowodowym zapisie pozwala na podjęcie próby rekonstrukcji niezrozumiałych wypowiedzi. Należy stanowczo zaznaczyć, że optymalnie wykonana korekcja nagrania powinna być kompromisem pomiędzy możliwie maksymalną redukcją zakłóceń a minimalną ingerencją w strukturę sygnału użytecznego niepowodującą jego deformacji. Łatwo wywnioskować, że jest to często trudne zadanie, nawet w dobie zaawansowanych technik do cyfrowego przetwarzania sygnałów.

Najważniejsze pojęcia związane z korekcją nagrań

Konieczne wydaje się sprecyzowanie najważniejszych pojęć i terminów, jakich używa się w związku z redukcją zakłóceń i korekcją nagrań dźwiękowych.

Pod względem praktycznym sygnałem użytecznym można nazwać taki sygnał, który zawiera istotne dla eksperta informacje przeznaczone do dalszego przetwarzania i analizy. Z kolei zakłóceniem można nazwać taki sygnał lub grupę sygnałów, które ograniczają lub uniemożliwiają odczytanie informacji z danego sygnału użytecznego. W odniesieniu do badań fonoskopijnych istotne jest określenie, co jest sygnałem użytecznym, a co stanowi zakłócenie. Nie jest to, jak by się mogło wydawać, zagadnienie trywialne. W procesie odtwarzania i spisywania treści wypowiedzi sygnałem użytecznym będzie oczywiście mowa, zaś zakłóceniami sygnały utrudniające jej zrozumienie, które w trakcie korekcji będą niwelowane. Dla badań identyfikacyjnych mowa również będzie sygnałem użytecznym. Badania autentyczności nagrania przeprowadza się m.in. z uwzględnieniem ujawnionych śladów magnetycznych, charakterystycznych sygnałów związanych z rejestratorem lub dodanych w procesie nagrywania (w tym przydźwięku sieciowego i zakłóceń od telefonów komórkowych) lub też innych, pozwalających zweryfikować ciągłość i oryginalność nagrania [11–14]. W tym przypadku zakłóceniem może być sygnał mowy. Zniekształcenie to deformacja prawidłowego przebiegu sygnału, która może wynikać z ograniczeń technicznych rejestratora, zastosowanej metody zapisu lub nieodpowiedniej obróbki sygnału. Filtracja to proces modyfikacji charakterystyk częstotliwościowych za pomocą filtra, w wyniku którego pewne pasmo częstotliwościowe ulegają tłumieniu (pasmo zaporowe filtra), pozostałe zaś nie (pasmo przepustowe filtra). Stosunek sygnału do szumu SNR (*signal-to-noise ratio*) to iloraz mocy albo amplitudy sygnału użytecznego do szumu (zakłócenia) wyrażony w decybelach (dB).

Proces korekcji nagrania niskiej jakości wraz z poprawą sygnału użytecznego jest zwykle czynnością wieloetapową.

Do elementów tego procesu można zaliczyć: zmianę parametrów nagrania (przez np. zwiększenie częstotliwości próbkowania i dekompresję), wyeksponowanie sygnału użytecznego, zminimalizowanie zakłóceń i zniekształceń oraz zmianę dynamiki nagrania. Ponieważ dowodowe nagrania odznaczają się różnymi parametrami i jakością, należy każdorazowo dokonać ich oceny i zastosować odpowiednią korekcję. Trzeba zaznaczyć, że nagranie poddane korekcji będzie zmodyfikowane w stosunku do nagrania źródłowego, co powinno się uwzględnić podczas dalszych badań z wykorzystaniem nagrania po korekcji.

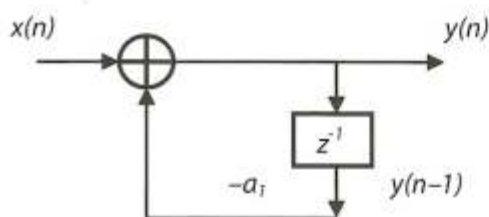
Metody redukcji zakłóceń występujących w dowodowych nagraniach dźwiękowych

Klasyczna filtracja sygnałów

Dzięki rozwojowi filtracji cyfrowej analiza sygnałów przeprowadzana jest obecnie w zdecydowanej większości za pomocą urządzeń i narzędzi cyfrowych. Dokonując ogólnej klasyfikacji filtrów cyfrowych, można wyszczególnić filtry rekursywne i nierekursywne. Te pierwsze cechują się nieskończoną odpowiedzią na impuls jednostkowy (*infinite impulse response – IIR*), która oscylując zmierza do zera, oraz odznaczają się tzw. sprzężeniem zwrotnym [15–16]. Oznaczając próbkę wejściową filtra jako $x(n)$, próbkę wyjściową filtra jako $y(n)$, zaś wagę jako a_1 , zasadę działania najprostszego filtra rekursywnego można zapisać wzorem:

$$y(n) = x(n) - a_1 y(n-1) \quad (1)$$

W wyniku filtracji podanym filtrem rekursywnym każda kolejna próbka $y(n)$ jest ważoną sumą aktualnej próbki $x(n)$ i poprzedniej próbki wyjściowej $y(n-1)$ [16], co przedstawiono na rycinie 1.

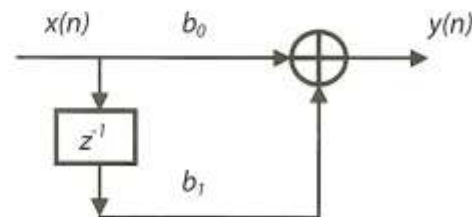


Ryc. 1. Struktura prostego filtra rekursywnego
Fig. 1. Structure of a simple recursive filter
Źródło (ryc. 1–10): autor

Filtry nierekursywne mają skończoną odpowiedź impulsową (*finite impulse response – FIR*), która po pobudzeniu funkcją impulsową zanika po odpowiednim czasie. Są to filtry bez sprzężenia zwrotnego. W najprostszym ujęciu filtra FIR sygnał wyjściowy $y(n)$ jest ważoną sumą dwóch próbek wejściowych, tj. obecnej $x(n)$ i poprzedniej $x(n-1)$ [15–16], co można przedstawić wzorem:

$$y(n) = b_0 x(n) + b_1 x(n-1) \quad (2)$$

Wagi b_0 i b_1 kształtują charakterystykę częstotliwościową filtra, określając pasmo zaporowe i przepustowe [16]. Na rycinie 2 przedstawiono strukturę opisanego filtra nierekursywnego.

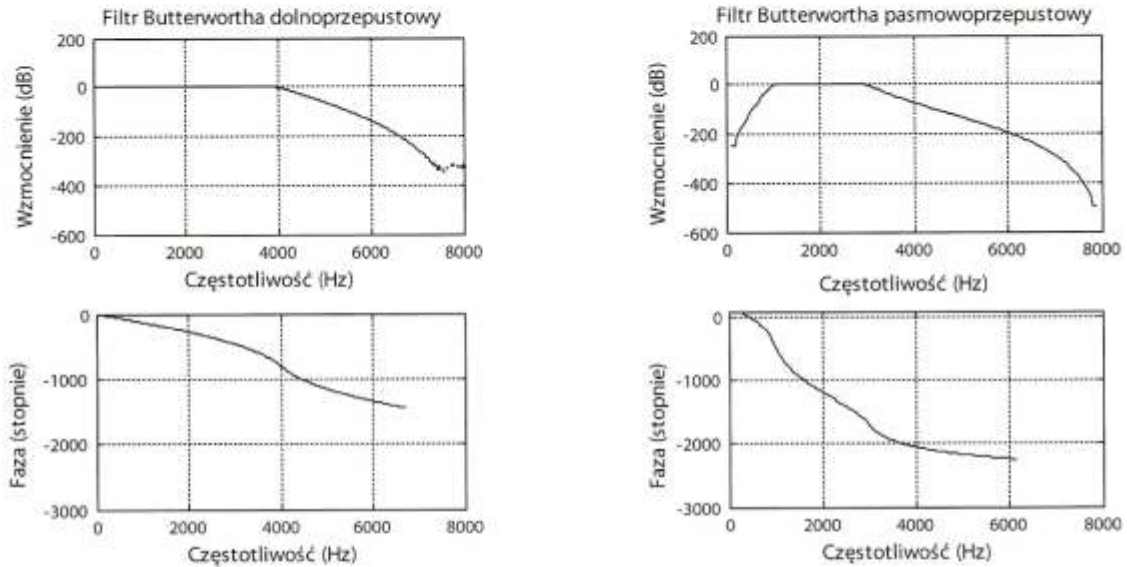


Ryc. 2. Struktura prostego filtra nierekursywnego
Fig. 2. Structure of a simple non-recursive filter

W celu uzyskania zamierzonych charakterystyk częstotliwościowych filtrów możliwe jest łączenie ze sobą prostych struktur filtrów lub obliczanie $y(n)$ na podstawie wielu próbek wejściowych lub wyjściowych i wielu wag. Dużą zaletą filtrów rekursywnych jest możliwość uzyskiwania stromych przejść między pasmami przepustowymi i zaporowymi przy niewielkiej liczbie wag. Wadą jest możliwość niestabilności filtra (wzbudzenie się) oraz nieliniowa charakterystyka fazowo-częstotliwościowa, co oznacza, że opóźnienie poszczególnych częstotliwości nie jest równe. Zaletą filtrów nierekursywnych jest liniowa charakterystyka fazowo-częstotliwościowa, która nie zniekształca sygnału, jak również stabilność filtra. Wadą jest duża złożoność obliczeniowa w stosunku do filtrów rekursywnych [15–16]. Do najbardziej popularnych w praktyce filtrów IIR należą: filtr Butterwortha, Chebyszewa (typ I i II), Bessela i eliptyczny.

Ze względu na kształty charakterystyk amplitudowo-częstotliwościowych filtry można podzielić na dolno-przepustowe, górnoprzepustowe, pasmowoprzepustowe i pasmowozaporowe. Na rycinie 3 przedstawiono charakterystyki amplitudowo-częstotliwościowe i fazowo-częstotliwościowe dla dwóch przykładowych filtrów Butterwortha, tj. dolnoprzepustowego i pasmowoprzepustowego.

Za pomocą opisanych wyżej filtrów można zredukować zakłócenia w nagraniach dźwiękowych w zdefiniowanych pasmach, przykładowo w wysokich lub niskich częstotliwościach, w których amplituda zostaje stłumiona. Pasmo zaporowe nie powinno obejmować zakresu częstotliwościowego, w którym znajduje się sygnał użyteczny. Szczególnym przykładem filtrów pasmowoprzepustowych i pasmowozaporowych są filtry określane, odpowiednio, *peak filter* i *notch filter*. Są to filtry o wąskim paśmie przepustowym albo zaporowym i bardzo stromym zboczem, służące do eksponowania albo tłumienia sygnałów o określonych częstotliwościach. Filtr *peak filter* można zastosować podczas badań autentyczności nagrań cyfrowych z wykorzystaniem sygnału przydźwięku, oscylującego wokół



Ryc. 3. Charakterystyki dwóch przykładowych filtrów Butterwortha: dolnoprzepustowy (po lewej, częstotliwość graniczna pasma przepustowego 4000 Hz) i pasmowoprzepustowy (po prawej, pasmo przepustowe od 1000 do 3000 Hz). W przypadku obu filtrów IIR widoczna jest nieliniowa charakterystyka fazowa
Fig. 3. Characteristics of two Butterworth filters: low-pass (on the left, cut-off frequency at 4000 Hz) and band-pass (on the right, pass band between 1000 and 3000 Hz). In the case of both IIR filters a nonlinear phase characteristic is visible

częstotliwości 50 Hz. Pasma przepustowe takiego filtra zawiera się w granicach około 49–51 Hz [12–14]. Filtry typu *notch filter* stosuje się do redukcji amplitudy sygnałów zakłócających o stałej lub niewiele zmieniającej się w czasie częstotliwości oraz ich harmonicznym. Są one przydatne, gdy częstotliwość zakłóceń znajduje się w obrębie pasma sygnału użytecznego.

Adaptacyjna filtracja sygnałów

Filtry typu FIR i IIR, których cechy przedstawiono w poprzednim podrozdziale, odznaczają się niezmiennością swoich parametrów po ich zaprojektowaniu i nie adaptują się do zmiany widma zakłócenia. W przypadku zakłóceń, których charakterystyka częstotliwościowa zmienia się w czasie, należy stosować filtry adaptacyjne [9, 10, 16, 17]. Na rycinie 4 przedstawiono strukturę przykładowego filtra adaptacyjnego, gdzie: $x(n)$ i $d(n)$ to sygnał filtrowany i odniesienia, natomiast $y(n)$ i $e(n)$ to sygnał po filtracji i błęd.

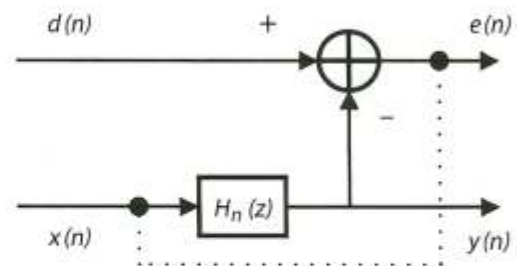
W procesie filtracji adaptacyjnej sygnał $x(n)$ zostaje przekształcony przez filtr opisany transmitancją $H_n(z)$. Współczynniki tego filtra zmieniają się w czasie (adaptują się) w taki sposób, aby sygnał $y(n)$ był jak najbardziej zbliżony do $d(n)$ [16]. Uwzględnia się przy tym założone kryterium błęd, np. średniokwadratowe, które można opisać wzorem:

$$J = E [e^2(n)] = E [(d(n) - y(n))^2] \quad (3),$$

gdzie $E \{ \cdot \}$ to wartość oczekiwana.

Filtry adaptacyjne mogą być filtrami IIR i FIR, jednakże częściej stosowane są rozwiązania oparte na filtrach FIR, z uwagi na złożoność algorytmów w przypadku IIR. Filtr adaptacyjny przedstawiony na rycinie 4 to filtr z tzw. kanałem referencyjnym (odniesienia). W praktyce nagrania dowodowe najczęściej rejestrowane są bez kanału odniesienia. W takich przypadkach stosuje się filtry adaptacyjne w dwóch możliwych konfiguracjach: detekcja sygnału użytecznego i usunięcie jego zakłóceń albo detekcja takiego sygnału i usunięcie go z nagrania. Ze względu na stosowane algorytmy wyróżnić można filtry adaptacyjne: LMS (*least mean square*), tzw. filtry bez pamięci, w których następuje minimalizacja chwilowej, nie zaś oczekiwanej wartości błęd, oraz filtry RLS (*root least square*), z pamięcią, w których miarą błęd jest kryterium najmniejszych kwadratów [16–17].

Na rycinie 5 zaprezentowano przykład zastosowania filtra adaptacyjnego typu FIR LMS do redukcji zakłócenia o zmiennych parametrach w czasie, tj. sinusoidy zmodulowanej częstotliwościowo.



Ryc. 4. Struktura przykładowego filtra adaptacyjnego
Fig. 4. Structure of a typical adaptive filter

Odejmowanie widma

Analizowane nagranie z zarejestrowanym badanym sygnałem, który został zakłócony, można opisać, wykorzystując założenie addytywności sygnałów. Zakłada się, że niezakłócony sygnał $x(t)$, np. mowa, zakłócony został sygnałem addytywnym $d(t)$, w wyniku czego otrzymano sygnał $y(t)$ zawierający sumę $x(t)$ i $d(t)$ [18]. Powyższe można przedstawić równaniem:

$$y(t) = x(t) + d(t) \quad (4).$$

Założenie to zostało zobrazowane na rycinie 6 z wykorzystaniem do tego celu reprezentacji czasowo-częstotliwościowej.

Przystępując do korekcji jako materiał początkowy mamy zazwyczaj takie nagranie, jakie zaprezentowano na rycinie 6c, w którym np. podczas rejestracji dodane zostały zakłócenia. Zadaniem eksperta jest wykonanie czynności odwrotnej do tej, jaką opisano wzorem (4) i przedstawiono na rycinie 6. Aby to osiągnąć, można wykonać odejmowanie widma zakłócenia od zakłóconego nagrania [19–20]. Proces ten przebiega w trzech podstawowych krokach, opisanych poniżej i zobrazowanych na rycinie 7:

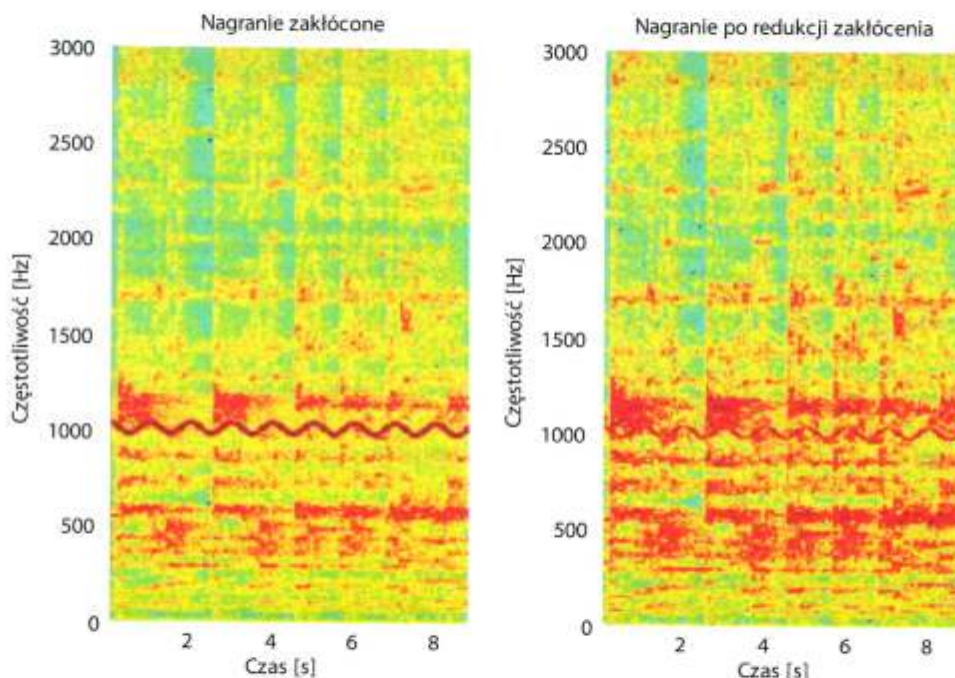
- wybór reprezentatywnego fragmentu nagrania zawierającego zakłócenie,
- wyznaczenie charakterystyki częstotliwościowej (widma) wybranego fragmentu,

- w dziedzinie częstotliwości wykonywanie operacji odejmowania charakterystyki widma wybranego fragmentu z zakłóceniem od kolejnych fragmentów zakłóconego nagrania.

Wykorzystując metodę redukcji zakłóceń przez odejmowanie widma, należy wziąć pod uwagę fakt, że z zakłóconego nagrania odjęte zostaną komponenty częstotliwościowe zawarte we wzorcu zakłócenia, tj. w kroku a). W związku z tym reprezentatywny fragment zakłócenia nie powinien obejmować sygnału użytecznego. Zaletą tej metody jest jej nieskomplikowany algorytm działania. Do wad zaliczyć można brak adaptacji do zakłóceń o zmiennej charakterystyce oraz powstawanie tzw. szumu muzycznego, który ujawnia się przy zbyt intensywnej redukcji zakłóceń w celu uzyskania jak największej wartości SNR.

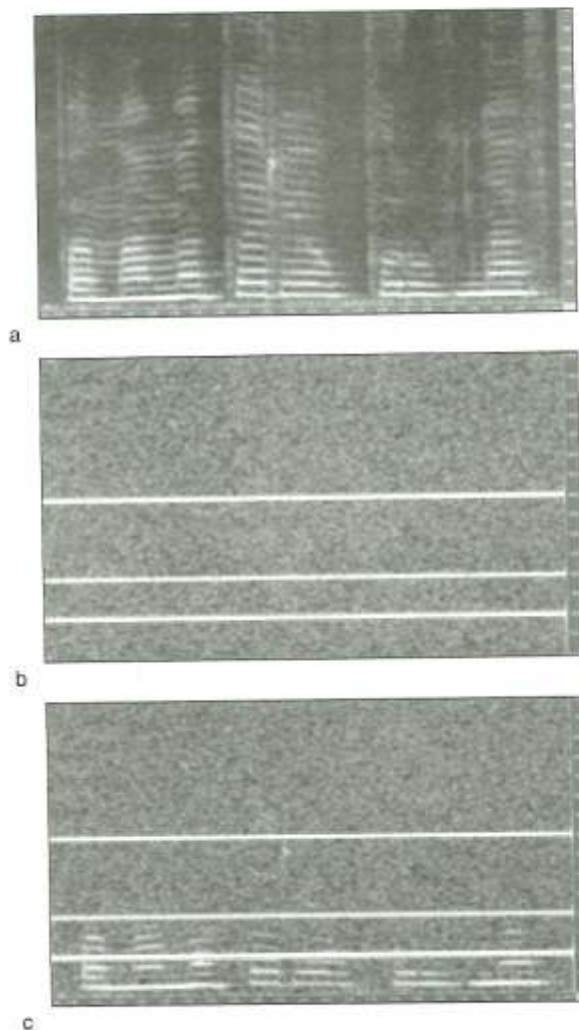
Redukcja zakłóceń wraz z estymacją sygnału mowy

Metody redukcji zakłóceń – klasyczna filtracja sygnałów i odejmowanie widma – mają tę istotną niedogodność, że nie są w stanie zaadaptować się do zmiennych w czasie charakterystyk zakłóceń, często występujących w nagraniach dowodowych. W takim przypadku możliwe jest zastosowanie tych metod do fragmentów nagrania, w których parametry zakłóceń są stałe lub niewiele zmieniające się i zaprojektowanie indywidualnych filtrów dla każdego fragmentu.



Ryc. 5. Nagranie zakłócone zmodulowaną częstotliwościowo sinusoidą 1000 Hz (po lewej) oraz to samo nagranie po redukcji zakłócenia za pomocą filtra adaptacyjnego typu FIR LMS (po prawej)

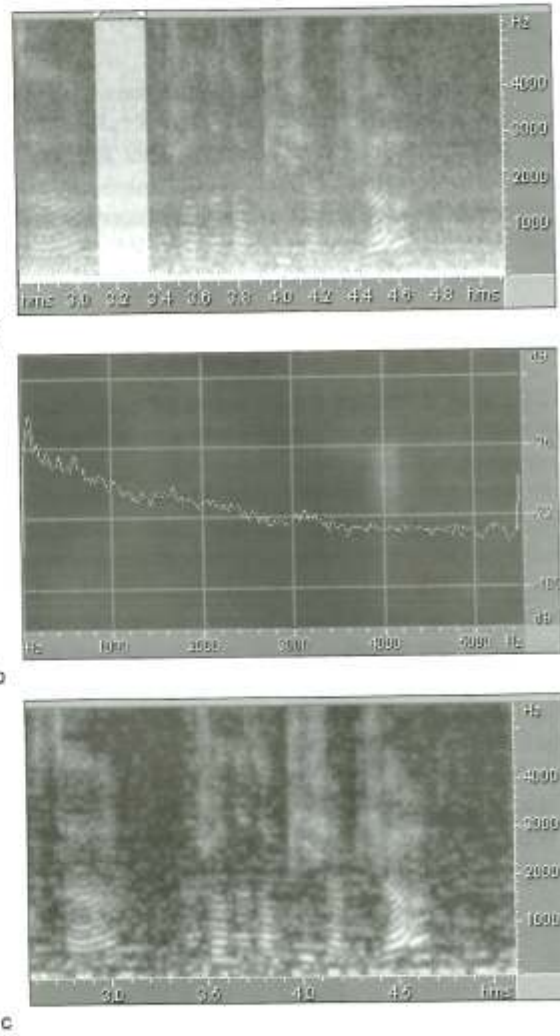
Fig. 5. Noisy recording with frequency modulated 1000 Hz sin-wave (on the left) and this recording after noise reduction with use of FIR LMS adaptive filter (on the right)



Ryc. 6. Zobrazowanie założenia addytywności sygnałów w nagraniu: a) spektrogram nagrania zawierającego niezakłócony sygnał mowy $x(t)$, b) spektrogram przykładowych zakłóceń $d(t)$ (szum biały i trzy przebiegi sinusoidalne), c) spektrogram sygnału mowy z dodanymi zakłóceniami $y(t)$. Osie odciętych reprezentują czas, zaś osie rzędnych – częstotliwość
Fig. 6. Visualization of the assumption of signals additivity in the recording: a) spectrogram of the recording with clean speech signal $x(t)$, b) spectrogram of noise $d(t)$, c) spectrogram of speech signal with added noise $y(t)$. Horizontal axes represent time and vertical axes – frequency

Omówiona wcześniej adaptacyjna filtracja sygnałów przydatna jest w praktyce do redukcji zakłóceń o charakterze okresowym lub prawie okresowym. Do redukcji zakłóceń zmieniających swoje parametry w czasie, głównie szerokopasmowych, dobrym rozwiązaniem są metody estymacji amplitudy sygnału użytecznego. Przykładem jest estymacja amplitudy sygnału mowy w krótkookresowej analizie spektralnej z minimalizacją błędu średniokwadratowego (*minimum mean square error short time spectral amplitude estimator – MMSE STSA estimator*) [9, 10, 18].

Niech $X_n(k) = |X_n(k)| \exp(j\alpha_n)$, $Y_n(k) = |Y_n(k)| \exp(j\theta_n)$ oraz $D_n(k)$ będą dyskretnymi transformatami Fourierowskimi, odpowiednio niezakłóconego i zakłóconego sygnału mowy oraz zakłócenia, n to numer analizowanej ramki sygnału,



Ryc. 7. Redukcja zakłóceń w nagraniach dźwiękowych metodą odejmowania widma: a) wybór reprezentatywnego fragmentu z zakłóceniem, b) charakterystyka widmowa fragmentu, c) spektrogram nagrania po redukcji zakłóceń
Fig. 7. Noise reduction in audio recordings with use of spectral subtraction method: a) selection of representative fragment of noisy recording, b) frequency characteristics of selected fragment, c) spectrogram of the recording after noise reduction

a k to indeks w dziedzinie częstotliwości. Uwzględniając powyższe, można wyznaczyć wartość *a priori* SNR na podstawie wzoru:

$$\epsilon_n(k) = \frac{E\{|X_n(k)|^2\}}{E\{|D_n(k)|^2\}} \quad (5)$$

oraz wartość *a posteriori* SNR według zależności:

$$\gamma_n(k) = \frac{|Y_n(k)|^2}{E\{|D_n(k)|^2\}} \quad (6),$$

gdzie $E\{\cdot\}$ to wartość oczekiwana. Za pomocą *a priori* SNR i *a posteriori* SNR możliwe jest wyznaczenie wartości

wzmocnienia $G_n(k)$. Estymacja amplitudy zrekonstruowanego sygnału mowy $x(t)$ wyznaczana jest przez przemnożenie wartości $G_n(k)$ z funkcją korelacji $R_n(k)$ [18].

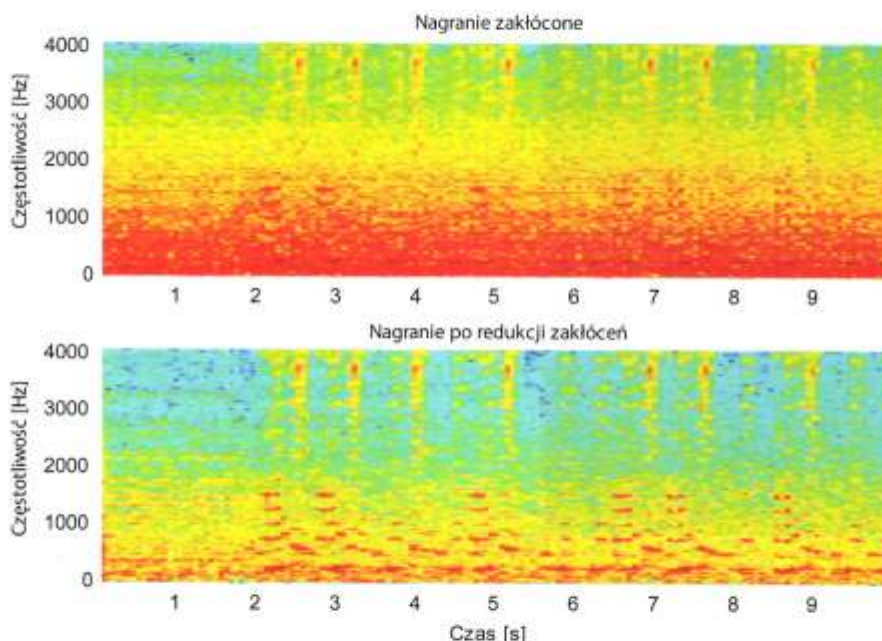
Przedstawiony algorytm pozwala na efektywne zwiększenie wartości stosunku SNR przez redukcję zakłóceń i estymację amplitudy sygnału mowy. Nagranie, które poddane zostaje procesowi redukcji zakłóceń, dzielone jest na mniejsze fragmenty (ramki), w każdym wykonuje się przedstawione wyżej obliczenia, a parametry algorytmu dostosowywane są adaptacyjnie w przypadku zmiany zakłócenia. Badania wykazały, że algorytm MMSE STSA odznacza się dobrą redukcją zakłóceń i poprawą stosunku SNR przy niewielkim nakładzie obliczeniowym, nieznacznych zniekształceniach oraz uwzględnieniu charakterystyki percepcji dźwięków przez człowieka [10]. Istotne jest również, że można go zastosować do nagrań jednokanałowych, czyli takich jak większość nagrań dowodowych. Na rycinie 8 przedstawiono wynik zastosowania algorytmu MMSE STSA zaimplementowanego w środowisku MATLAB do redukcji zakłóceń z nagrania dźwiękowego [10, 21].

Redukcja zakłóceń i zniekształceń w dziedzinie czasu

Przedstawione powyżej metody redukcji zakłóceń odnosiły się do analizy sygnałów w dziedzinie częstotliwości. Analiza częstotliwościowa jest to podstawowy etap wykonywany podczas korekcji nagrań, w tym również do-

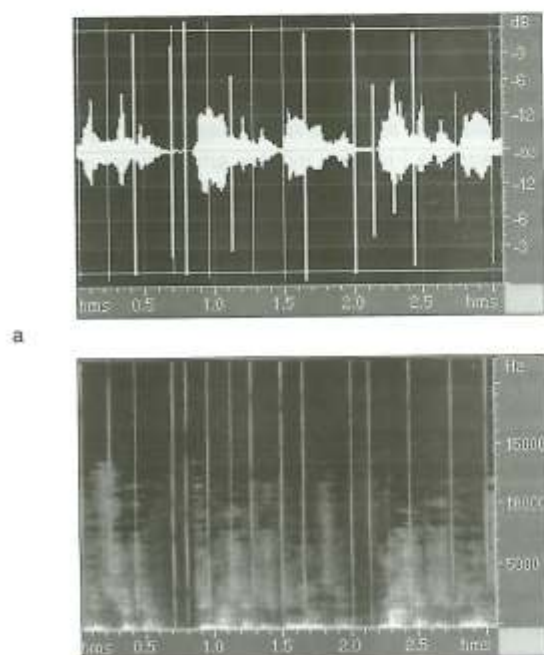
wodowych. Redukcja zakłóceń możliwa jest także przez modyfikację przebiegu czasowego nagrania. Jedną z często stosowanych technik podczas korekcji jest kompresja dynamiki. Ma ona za zadanie zmniejszenie zbyt dużej dynamiki nagrania i różnicy między amplitudami największymi a najmniejszymi [19]. Redukcję zbyt dużej amplitudy sygnału można przeprowadzić z zastosowaniem algorytmu limitera, który ogranicza amplitudę do zdefiniowanego maksymalnego poziomu. Za pomocą kompresji dynamiki i limitera możliwe jest minimalizowanie zakłóceń o charakterze impulsowym, tj. krótkotrwałych i zazwyczaj szerokopasmowych pików amplitudowych [9, 20]. Należy dodać, że do tego typu zakłóceń stosowane są również dedykowane algorytmy, analizujące zmiany amplitudy przebiegu (redukcja klików i trzasków). Podczas stosowania opisanych technik należy dobrać odpowiedni poziom redukcji, aby, zwłaszcza przy dużych ograniczeniach, uniknąć niepożądanych zniekształceń sygnału [19]. Na rycinie 9 zaprezentowano przykładowe nagranie zawierające zakłócenia impulsowe.

W przypadku, gdy amplituda sygnału nagrywanego przekroczy maksymalną wartość akceptowalną dla urządzenia rejestrującego, często występuje efekt tzw. przesterowania. Objawia się on zniekształceniami przebiegu czasowego sygnału, głównie wysokich amplitud. Na rycinie 10 zaprezentowano fragment nagrania z widocznym efektem przesterowania, a także ten sam fragment po redukcji tego typu zniekształceń. Minimalizację przesterowania osiągnięto w wyniku zastosowania kompresji



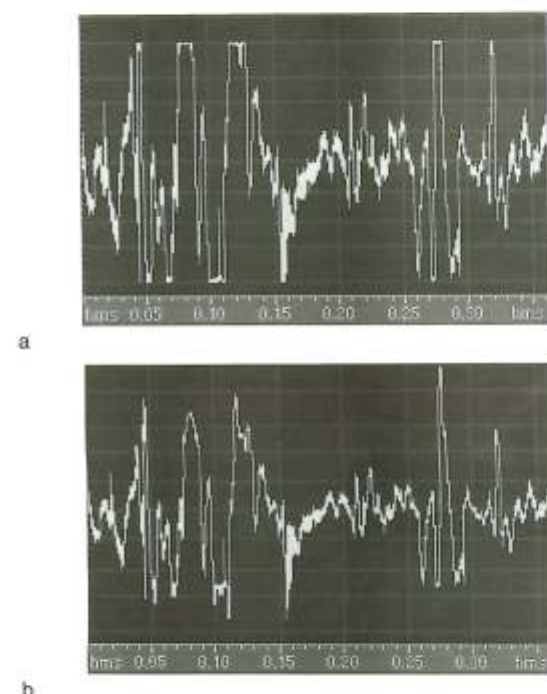
Ryc. 8. Spektrogram nagrania dźwiękowego przed (na górze) i po redukcji zakłóceń (na dole) za pomocą algorytmu MMSE STSA. Wykorzystując tę technikę, można efektywnie redukować zakłócenia nawet w nagraniach o bardzo małym SNR, jak na niniejszej rycinie (u góry)

Fig. 8. Spectrogram of the recording before (on the top) and after noise reduction (at the bottom) with use of MMSE STSA algorithm. This technique enables effectively reduction of noise even if the recording is characterized with very low SNR, as shown (on the top)



Ryc. 9. Charakterystyka amplitudowa (a) oraz spektrogram (b) nagrania z utrwalonymi zakłóceniami o charakterze impulsowym

Fig. 9. Amplitude characteristics (a) and spectrogram (b) of the recording with impulse noises



Ryc. 10. Charakterystyka fragmentu nagrania z efektem przesterowania (a) objawiającym się zniekształceniem przebiegu czasowego (widoczne „obcięcie” amplitudy sygnału) oraz charakterystyka fragmentu (b) po zastosowaniu algorytmu do redukcji tego efektu (rekonstrukcja zniekształconego przebiegu)

Fig. 10. Characteristics of the fragment of the recording with overload signal phenomenon (a), which can be observed as distortion of the wave (clipped amplitude) and characteristics of this fragment (b) after applying the algorithm for reduction of such phenomenon (reconstruction of distorted signal)

dynamiki, którą opisaną powyżej, jak również algorytmu wykorzystującego tzw. funkcje sklepane (*spline functions*). W algorytmie tym „skleja” się ze sobą funkcje wielomianów niskiego stopnia (na przykład drugiego albo trzeciego), co prowadzi do rekonstrukcji zniekształconego przebiegu [15, 21].

Podsumowanie

W niniejszej pracy dokonano przeglądu i zestawienia wybranych metod wykorzystywanych do redukcji zakłóceń, jakie obserwowane są często w dowodowych nagraniach dźwiękowych. Jak zaznaczono, jednym z podstawowych badań fonoskopijnych jest analiza audytywna nagrania i spisanie treści wypowiedzi, co stanowiło główną wytyczną podczas wyboru prezentowanych metod. Obecnie dostępnych jest wiele narzędzi przeznaczonych do cyfrowego przetwarzania sygnałów. W niniejszej pracy przedstawiono wybrane metody redukcji zakłóceń, które stosowane są w praktyce. W celu poprawy jakości sygnału mowy wykonywana jest korekcja nagrania, której zadaniem jest maksymalizacja odczytu informacji z sygnału użytecznego przy zminimalizowaniu maskujących go zakłóceń i zniekształceń. Z uwagi na to, że zastosowanie każdego narzędzia do redukcji zakłóceń jest ingerencją w zapis sygnału dowodowego, należy stosować jak najmniejszą ich liczbę, wystarczającą do otrzymania zadawalających rezultatów. Ponieważ dowodowe nagrania odznaczają się różnymi parametrami oraz różną zawartością zakłóceń i zniekształceń, nie istnieje korekcja uniwersalna. Wykonanie jej wymaga indywidualnej analizy każdego nagrania oraz zastosowania przez eksperta odpowiednich metod i narzędzi.

BIBLIOGRAFIA

1. Błasikiewicz S.: Metoda odsłuchu szeptu i mowy intensywnie zakłóconej, „Problemy Kryminalistyki” 1971, nr 90, s. 159–183.
2. Błasikiewicz S., Miściuk A., Wójcik W.: Podstawowy zakres badań fonoskopijnych prowadzonych w Zakładzie Kryminalistyki KG MO, „Problemy kryminalistyki” 1967, nr 67–68, s. 303–327.
3. Malanowicz A., Koźmiński L.: Fonoskopia, Wydawnictwo Szkoły Policji w Pile, Piła 2009.
4. Rzeszotarski J.: Kompendium badań fonoskopijnych, „Prokuratura i Prawo” 2007, nr 7–8, s. 181–187.
5. Suszczewski W.: Ekspertyza fonoskopijna, [w:] Ekspertyza sądowa. Zagadnienia wybrane, Wójcikiewicz J. [red.], Oficyna a Wolters Kluwer business, Warszawa 2007.
6. Iwanicka B., Koprowski E.: Kasety magnetofonowe i magnetowidowe, Wydawnictwa Komunikacji i Łączności, Warszawa 1988.

7. Sztekmiler K.: Podstawy nagłośnienia i realizacji nagrań. Podręcznik dla akustyków, Narodowe Centrum Kultury, Warszawa 2003.

8. Korycki R.: Wielokanałowe techniki korekcji nagrań fonicznych w kryminalistyce, „Problemy Kryminalistyki” 2011, nr 273.

9. Michałek M.: Noise reduction in evidence audio recordings with use of speech spectral amplitude estimation, ENFSI Working Group for Forensic Speech and Audio Analysis (FSAAWG), Rome, Italy, 15–16 September 2011.

10. Michałek M.: Noise reduction in audio recordings for forensic audio research using an algorithm for short-time spectral analysis, „Problems of Forensic Science” 2013 (akceptacja do druku).

11. Dean D.J.: The Relevance of Replay Transients in the Forensic Examination of Analogue Magnetic Tape Recordings, Scientific Research and Development Branch, Home Office, British Government, London 1991, no. 16.

12. Grigoras C., Cooper A., Michałek M.: Forensic Speech and Audio Analysis Working Group – Best Practice Guidelines for ENF Analysis in Forensic Authentication of Digital Evidence, ENFSI – FSAAWG 2009.

13. Michałek M.: The application of powerline hum in digital recording authenticity analysis, „Problems of Forensic Sciences” 2009, vol. 80, p. 355–364.

14. Michałek M.: Automatyzacja badania autentyczności cyfrowych nagrań dźwiękowych z zastosowaniem kryterium ENF, „Problemy Kryminalistyki” 2010, nr 267, s. 26–32.

15. Uhl T., Kurowski P.: Zastosowanie środowisk MATLAB i Siglab do analizy sygnałów, Tom 13 z CCATIE Cracow Centre for Advanced Training in Information Engineering, Katedra Automatyki AGH, 1998.

16. Zieliński Tomasz P.: Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań, Wydawnictwa Komunikacji i Łączności, Warszawa 2007.

17. Kustra G.: Algorytmy identyfikacji i adaptacji w jednokanałowych systemach aktywnej redukcji hałasu, Materiały IX Sympozjum Naukowego „Nowości w Technice Audio i Wideo”, Warszawa 2002, s. 123–135.

18. Ephraim Y., Malah D.: Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator, IEEE Transactions on Acoustic, „Speech and Signal Processing” 1984, vol. 32, p. 1109–1121.

19. Korbecki M.: Komputerowe przetwarzanie dźwięku, Wydawnictwo MIKOM, Warszawa 1999.

20. Smoliński R.: Cyfrowe systemy rekonstrukcji dźwięku, Materiały VII Sympozjum Naukowego „Nowości w Technice Audio”, Warszawa 2000, s. 141–154.

21. MATLAB User’s Guide R2012b.

kryminalistycznych. Spisanie treści wypowiedzi z nagrań jest przedmiotem niemal każdej ekspertyzy fonoskopijnej. W dowodowych nagraniach dźwiękowych bardzo często występują liczne zakłócenia, charakteryzujące się dużą amplitudą oraz zakresem częstotliwościowym skutecznie maskującym sygnał mowy. Zakłócenia te mogą wynikać ze sposobu nagrywania, z warunków akustycznych podczas zdarzenia jak i użytej techniki rejestracji. Wiąże się to z koniecznością wykonywania korekcji nagrań w celu poprawy jakości sygnału mowy. W artykule opisano najważniejsze terminy odnoszące się do redukcji zakłóceń w nagraniach dźwiękowych oraz zastosowanie filtrów cyfrowych typu IIR, FIR oraz selektywnych peak i notch. Oprócz klasycznej filtracji sygnałów duże znaczenie w tej dziedzinie mają metody do adaptacyjnej redukcji zakłóceń. Praca przedstawia również właściwości filtrów adaptacyjnych LMS i RLS, algorytmu do estymacji amplitudy sygnału mowy w oparciu o estymator MMSE STSA, jak również metodę odejmowania widma. Oprócz technik wykorzystujących analizę częstotliwościową w artykule zaprezentowano sposoby redukcji zakłóceń w dziedzinie czasu. Wybrane metody zostały opisane teoretycznie. Artykuł zawiera także przykłady praktycznego zastosowania zaprezentowanych filtrów i algorytmów.

Słowa kluczowe: fonoskopia, filtracja sygnałów, redukcja zakłóceń, korekcja nagrania

Summary

The aim of the paper is to present the issue of noise reduction in audio recordings for forensic research. Speech to text transcription is a subject of almost every forensic audio analysis. In evidence audio recordings quite often there are numerous noises characterized by a high volume and frequency which effectively mask the speech signal. These noises may result from the way of recording, acoustic conditions during the event and the technique of registration used. It is connected with the necessity to make correction of recordings to improve the quality of speech signal. The paper describes the most important technical terms related to noise reduction in audio recordings and the use of digital filters like IIR, FIR, peak and notch. In addition to classical signal filtering methods adaptive noise reduction algorithms play an important part in this case. The article presents some features of LMS and RLS adaptive filters, algorithm for speech amplitude estimation based on MMSE STSA estimator and spectral subtraction method as well. Apart from methods using frequency analysis the article contains some details of noise reduction in time domain. Selected methods were described theoretically. The article also contains examples of practical application of presented filters and algorithms.

Keywords: forensic audio analysis, signal filtering, noise reduction, improving recording quality.

Streszczenie

Celem niniejszego artykułu jest przedstawienie problematyki redukcji zakłóceń w nagraniach dźwiękowych z uwzględnieniem badań