

RZECZPOSPOLITA
POLSKA



Urząd Patentowy
Rzeczypospolitej Polskiej

(12) **OPIS PATENTOWY**

(19) **PL**

(11) **223653**

(13) **B1**

(21) Numer zgłoszenia: **402147**

(51) Int.Cl.
G10K 11/00 (2006.01)

(22) Data zgłoszenia: **20.12.2012**

(54)

Sposób eliminacji przesłuchów

(43) Zgłoszenie ogłoszono:

23.06.2014 BUP 13/14

(45) O udzieleniu patentu ogłoszono:

31.10.2016 WUP 10/16

(73) Uprawniony z patentu:

**AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE,
Kraków, PL**

(72) Twórca(y) wynalazku:

ROBERT BARAŃSKI, Kraków, PL

PL 223653 B1

Opis wynalazku

Wynalazek dotyczy sposobu filtracji przesłuchów. Sposób ten może być wykorzystywany w sytuacjach występowania przesłuchów zarówno w nagraniach akustycznych jak i w innych sytuacjach kiedy mamy do czynienia z powtarzającymi się sygnałami np. w postaci impulsów.

Eliminacja przesłuchów, zarówno w nagraniach muzycznych, jak i w innych sytuacjach jest zagadnieniem, które niejednokrotnie było tematem zainteresowań i publikacji.

W studiach nagraniowych przy wykorzystywaniu technik wielomikrofonowych występuje zjawisko przesłuchu. Każdy mikrofon z założenia powinien rejestrować sygnał tylko z konkretnego instrumentu mu przypisanego (np. werbla lub kotła). Jednak z uwagi na zjawisko rozchodzenia dźwięku, część energii akustycznej dociera do pozostałych mikrofonów. W związku z powyższym poza swoim docelowym instrumentem rejestrują one pozostałe instrumenty. Tak więc rejestracja innych instrumentów przez mikrofon przeznaczony do rejestracji konkretnego instrumentu identyfikuje się jako przesłuch.

Na stronie internetowej <http://www.ambiophonics.org/PCMac.html> przedstawiono przykład aplikacji, która eliminuje przesłuchy przy wykorzystaniu algorytmu RACE (Recursive Ambiophonic Cross-talk Elimination). Jednak jego zastosowanie nie jest przewidziane do eliminacji przesłuchów powstałych podczas nagrań muzycznych.

Na stronie internetowej <http://dl.acm.org/citation.cfm?id=981075> przedstawiono przykład zastosowania algorytmu do eliminacji szeroko pojętych przesłuchów. W tym przypadku publikacja związana jest z produkcją półprzewodników. Jednak opisany algorytm nie jest wykorzystywany do eliminacji przesłuchów powstałych podczas nagrań muzycznych. Nie opiera się on również na zasadzie wielopakietowej filtracji falkowej.

Znane jest ze strony internetowej <http://www.cstr.ed.ac.uk/downloads/publications/2005/sap04-xtalk.pdf> rozwiązanie, w którym dokonuje się eliminacji przesłuchów opierając się na korelacjach (m.in. auto correlation, cross correlation). System niezbędny do analizy wymagał rejestracji sygnału rejestrowanego dwoma mikrofonami a więc operacje są wykonywane z użyciem co najmniej dwóch ścieżek nagraniowych.

W sposobie eliminacji przesłuchów przy realizacji nagrań dźwiękowych według wynalazku za pomocą pierwszego urządzenia rejestrującego sygnał rejestruje się sygnał z pierwszego urządzenia źródła sygnału, jednocześnie w różnych odstępach czasu przez pierwsze urządzenie rejestrujące sygnał rejestruje się niechciane sygnały z drugiego źródła sygnału, następnie dla wybranego wzorcowego impulsu wykonuje się wielopakietową transformatę falkową sygnału, w skutek czego otrzymuje się wzorcowe wartości kryterialne analizowanego fragmentu dla każdej ze skal, czyli dla każdego z pasm częstotliwości, a następnie wielopakietowej transformacie falkowej poddaje się wszystkie pozostałe sygnały z przesłuchem, po czym porównuje się fragmenty sygnału z przesłuchem z odpowiednimi fragmentami sygnału wzorcowego, przy czym gdy fragmenty sygnału z przesłuchem znajduje się poza z góry określonym obszarem następuje zamiana fragmentu sygnału z przesłuchem na fragment sygnału bez przesłuchu.

Wielopakietową transformatę falkową można wykonywać w środowisku programowym takim jak np. Matlab lub LabView. Operacje można też wykonywać na odpowiednio przygotowanych układach FPGA.

Celem wynalazku jest dostarczenie sposobu eliminacji przesłuchów przy realizacji nagrań dźwiękowych.

Wynalazek ten umożliwia skuteczne i szybkie eliminowanie przesłuchów. A w przypadku odpowiedniej implementacji również bardzo proste w obsłudze, ograniczające się wyłącznie do wskazania impulsu wzorcowego.

Wynalazek poza podstawową funkcją eliminacji przesłuchów posiada niezmiernie istotną własność przy nagraniach. W dobie automatyzacji, coraz częściej zwraca się uwagę na fakt, by nagranie brzmiało „żywo”. Sprowadza się to do tego, by np. następujące po sobie uderzenia werbla nie były identyczne w równych odstępach (jak wygenerowane przez automat). Proponowany algorytm w nagraniu ingeruje wyłącznie w przesłuch. Oznacza to, iż czas występowania sygnału (np. uderzenia werbla), jego charakter brzmieniowy oraz amplituda (np. siła uderzenia w werbel) pozostają indywidualne. Ich charakter nie jest zatracany.

Obecnie znane rozwiązania często opierają się na zamianie całego fragmentu nagrania, całkowicie zmieniając jego charakter.

Przedmiot wynalazku w przykładzie wykonania jest uwidocznił na rysunku, na którym Fig. 1 przedstawia schemat działań przeprowadzanych w sposobie według wynalazku;

Fig. 2 przedstawia wyniki zebranych próbek sygnału wzorcowego AS_i, sygnału z przesłuchem ABS_i i sygnału uzyskanego po przeprowadzeniu sposobu według wynalazku;

Fig. 3 przedstawia wyniki przebiegów czasowych dla filtracji falką Daubechies 2, stopień dekompozycji d=3;

Fig. 4 przedstawia schemat jak wraz ze zwiększaniem się stopnia dekompozycji rośnie ilość pasm.

Sposób według wynalazku opiera się na podejściu do nagrania jako bloków, w których rozpatrywany fragment posiada stały charakter. W takich blokach, wykonywana jest analiza poszczególnych sygnałów impulsowych w celu znalezienia „wzorca” sygnału. Na jego podstawie budowana jest baza filtrów (przy wykorzystaniu wielopakietowej dekompozycji falkowej). Baza ta następnie jest wykorzystywana do rekonstrukcji pasm dla impulsów (uderzeń), których charakter jest mocno odmienny od wzorca.

W ten sposób, nie tracąc w pełni indywidualnego charakteru każdego z impulsów (uderzeń), niwelowane są różnice pomiędzy sąsiadującymi impulsami (uderzeniami), przez co charakter nagrania staje się bardziej jednorodny (w rozpatrywanych fragmentach np. zwrotce). Narzędzie to może okazać się pomocne przy miksowaniu nagrań instrumentów perkusyjnych w przypadku, gdy dotyczy ono niedoświadczzonego perkusisty grającego nierówno w rozumieniu brzmienia oraz amplitudy (nie ingerując w czas występowania sygnału).

Na Fig. 1 przedstawiono w większych szczegółach kolejne kroki sposobu według wynalazku. Na początku, w etapie 1 rejestruje się sygnał. Warto nadmienić, iż algorytm jest dedykowany do pracy na sygnale z zarejestrowanymi przebiegami impulsowymi (pochodzącymi np. z nagrań perkusyjnych). Jako impuls rozumiany jest sygnał o krótkim czasie trwania, rozpoczynający się wysoką amplitudą zanikającą w czasie, przeważnie czas trwania to 1–5 sekund. Przykładem może być uderzenie w werbel.

Sposób realizacji nagrania nie jest istotny. Może to być nagranie wykonane z wykorzystaniem mikrofonu i dowolnego rejestratora (np. ZOOM H4n) lub komputera wyposażonego w kartę muzyczną (zintegrowaną lub dedykowaną jak np. ASUS X0NAR DG). Nagranie może również pochodzić z Internetu albo dowolnego nośnika danych np. płyty CD itp.

Jedynym stawianym wymaganiem – opisywany sposób pracuje na sygnałach jednokanałowych o znanej częstotliwości próbkowania.

Następnie przechodzimy do etapów 2–7, które są wykonywane dla każdego impulsu indywidualnie.

W etapie 2 dokonuje się filtracji WDP impulsu „i”. Dla wybranego przez użytkownika impulsu „i” (czyli fragmentu nagrania) wykonuje się wielopakietową transformatę falkową WPD (Wavelet Packet Decomposition). Jest to powszechnie znana transformata i nie jest ona przedmiotem niniejszego zgłoszenia, więc jej szczegółowy opis zostanie pominięty.

W wyniku działania WPD otrzymujemy ‘n’ przebiegów czasowych $f_n(t)$. Ich ilość zależy od wyboru pakietu filtrów falkowych (w etapie 9), a wprost $n=2^d$, gdzie d – stopień dekompozycji. Przykładowe wyniki przebiegów czasowych dla filtracji falką Daubechies 2, stopień dekompozycji d=3 przedstawione są na Fig. 3.

W etapie 3 oblicza się wartość entropii $E_i(n)$ dla impulsu „i”, przy czym n oznacza numer pasma zależny od stopnia dekompozycji, przykładowo $n=2^d$.

Ilość wykorzystywanych pasm częstotliwości określa liczba ‘d’. Przykładowo przy próbkowaniu wynoszącym 2000 Hz mamy do czynienia z częstotliwością użyteczną wynoszącą 1000 Hz (wynika to z kryterium Nyquista). Dla stopnia dekompozycji wynoszącego d=2, czyli gdy liczba pasm wynosi n=4, otrzymuje się pasma o szerokości wynoszącej 250 Hz, a więc 0–250 Hz, 250–500 Hz itd. Natomiast dla stopnia dekompozycji wynoszącego d=4, czyli gdy liczba pasm wynosi n=16 otrzymujemy pasma o szerokości 1000/16=62.5 Hz, a więc 0–62.5 Hz, 62.5–125 Hz itd.

Obliczenie wartości dla każdego z otrzymanych przebiegów czasowych $f_k(t)$, będzie stanowiło element kryterium sposobu. Przewidziano użycie takich parametrów jak:

- Entropia $Entropia_n = \sum \ln(x^2)$.
- Energia $Energia_n = \sum x^2$
- Wartość skuteczna (RMS) $RMS_n = \sqrt{\frac{\sum(x^2)}{n}}$

Następnie przechodzimy do etapu 4, w którym sprawdza się nierówności

$$E_i(n) < (1-k)E_w(n) \quad \text{lub} \quad E_i(n) > (1+k)E_w(n)$$

Przy czym $E_w(n)$ oznacza wartość wybranego parametru (np. entropię) sygnału wzorcowego w paśmie n , a $E_i(n)$ wartość wybranego parametru (np. entropię) sygnału badanego w paśmie n .

Jeśli powyższe nierówności są spełnione dokonuje się zmiany sygnału badanego na wzorcowy w paśmie n , w przeciwnym przypadku badany sygnał pozostawia się bez zmian.

Warunek sprawdzający czy wartość parametru wybranego w etapie 3 (np. entropia) nie różni się bardziej niż o wartość k (wyrażona w procentach).

Jeśli różni się, wykonywany jest etap 5, w przeciwnym przypadku wykonywany jest etap 6.

Po przejściu do etapu 5 usuwa się przebieg czasowy pasma 'n' analizowanego impulsu 'i' (fragmentu sygnału), wstawiając w jego miejsce przebieg czasowy sygnału wzorcowego odpowiedniego dla pasma 'n'.

W etapie 6 nie wykonuje się żadnych operacji i pozostawia się sygnał bez zmian.

Etapy 3, 4, 5 i 6 wykonuje się 'n' razy (czyli tyle ile jest pasm powstałych w wyniku operacji WPD).

Potem przechodzi się do etapu 7, w którym dokonuje się odwrotnej transformaty falkowej.

Z uwagi na fakt, iż jest to ogólnie znana transformata, oraz nie jest ona przedmiotem niniejszego zgłoszenia to jej szczegółowy opis zostanie pominięty.

Jej zadaniem jest synteza wszystkich 'n' przebiegów w jeden przebieg. Wynikiem tej operacji jest jeden, sumaryczny przebieg czasowy będący złożeniem wszystkich 'n' przebiegów czasowych.

Jeśli podczas wykonywania algorytmu zawsze wynikiem etapu 4 będzie etap 6, to nie dojdzie do zamiany żadnego z analizowanych sygnałów, więc sygnał jaki potrzymamy na wyjściu będzie identyczny z tym co podano na wejście algorytmu.

W etapie 8 z wczytanego sygnału należy manualnie wybrać impuls (fragment sygnału), który będzie służył jako wzorzec. Wyboru dokonuje osoba obsługująca algorytm. Z reguły wybiera się impuls (fragment sygnału), który uważa się za najczystszy (bez przesłuchu), ładny brzmieniowo. Precyzyjniej, należy wskazać początek i koniec trwania impulsu.

W etapie 9 do realizacji wielopakietowej analizy falkowej WPD konieczne jest określenie dwóch parametrów: funkcji falkowej oraz stopnia dekompozycji.

W przypadku funkcji falkowych, znanych jest co najmniej kilkadziesiąt powszechnie używanych falek (czyli zestawów filtrów). Aby zestaw filtrów mógł zostać uznany za falekę musi spełniać pewne kryteria. Kryteria te są ogólnie znane, przykładowo może to być cała rodzina falek Daubechies (kilkanaście zestawów filtrów), Coiflets (kilkanaście zestawów filtrów). W związku z powyższym nie ma powodu by działanie algorytmu ograniczać do konkretnych falek, gdyż jest to po prostu filtracja. Różnicą jest zastosowanie filtrów zwanych falkowymi, czyli spełniających owe dodatkowe kryteria. Szczegół jaki to jest filtr nie jest w tym przypadku istotny. Natomiast jego dobór będzie miał wpływ na skuteczność filtracji. Niestety na obecnym etapie nie można podać dokładnie jakie filtry mają zostać zastosowane, gdyż ocena ich działania to subiektywne wrażenia użytkownika.

Jeśli chodzi o stopień dekompozycji, to z każdym stopniem dekompozycji ilość pasm na które dzieli się sygnał podwaja się. Tak więc ilość pasm 'n' odpowiada zależności $n=2^d$, gdzie d – stopień dekompozycji.

Fig. 4 przedstawia schemat jak wraz ze zwiększaniem się stopnia dekompozycji rośnie ilość pasm.

Jest to graficzna interpretacja wzoru dla $n=d^2$, gdzie A oznacza aproksymację, D to detal, i dalej AA2 to aproksymacja aproksymacji na poziomie nr 2, zaś AD2 to aproksymacja detalu na poziomie nr 2.

Zastosowane skróty wynikają z zasad oznaczania filtrów falkowych i wyników ich działania.

Mając wybrany pakiet filtrów falkowych o określonych parametrach dokonuje się filtracji falkowej impulsu (fragmentu sygnału) wzorcowego-etap 10. Operacja jest identyczna jak ta w etapie 2. Efektem tego działania jest uzyskanie 'n' przebiegów czasowych (dla każdego z n pasm częstotliwościowych).

Następnie, w etapie 11, oblicza się wartość wybranego parametru (np. entropii) dla sygnału wzorcowego $E_w(n)$ dla wszystkich 'n' pasm (skal).

Operacja identyczna jak w etapie 3. Ma ona na celu wyznaczenie wartości wybranego parametru (np. entropii). Efektem działania blocka jest uzyskanie co do ilości 'n' wartości (dla każdego pasma jedna). Będzie to wartość tzw. wzorcowa. Z nią porównywane są odpowiadające im wartości (np. entropii) obliczone dla impulsów w danym paśmie 'n'.

Dla ułatwienia prezentacji sposobu, przedstawiono konkretny przykład wykonania sposobu dla rejestracji perkusji.

Przykładowy zestaw perkusji zawiera werbel oraz kocioł. W czasie gry rejestruje się mikrofonem sygnał z urządzenia A, którym jest werbel. Są to impulsy odpowiadające uderzeniom w membranę werbla. Jednocześnie co jakąś chwilę czasową następuje zakłócenie z urządzenia B, którym jest kocioł. Na nagraniu otrzymujemy sygnał zawierający fragmenty bez przesłuchu (A) pochodzące tylko od werbla jak i z przesłuchem (AB) pochodzące od werbla i kotła. Podczas całego nagrania obserwujemy wielokrotne występowanie fragmentów pochodzących tylko od werbla przeplecionych fragmentami pochodzącymi od werbla i kotła.

W sposobie eliminacji według wynalazku przesłuchu wykonuje się wielopakietową transformatę falkową (WPD Wavelet Packet Decomposition) sygnału wzorcowego AS_n .

Po poddaniu sygnału AS transformacie WPD otrzymuje się informację jak brzmi czysty/idealny sygnał AS_n w obliczonych skalach (konkretnych pasmach częstotliwości) (AS_n – sygnał czysty przedstawiony w postaci skal, indeks n oznacza numer skali/pasma).

Następnie transformacie WPD poddaje się sygnał z przesłuchem AB uzyskując informację o brzmieniu sygnału w skalach (ABS_n – sygnał z zakłóceniami przedstawiony w postaci skal).

Kolejnym krokiem jest porównanie sygnałów z przesłuchem ABS_n z odpowiednimi sygnałami wzorcowymi AS_n . Kryterium porównania jest zapytanie czy różnica w np. energii sygnału z przesłuchem ABS_n oraz sygnału wzorcowego AS_n jest większa niż próg podawany w procentach względem sygnału AS_n . Czyli dla progu równego 10% warunek wygląda następująco:

$$\begin{aligned} &ABS_n \leq 90\%AS_n \\ &\text{lub} \\ &ABS_n \geq 110\%AS_n \end{aligned}$$

W przypadku gdy dla badanej skali (pasma) n warunek nie jest prawdziwy, następuje zamiana sygnału z przesłuchem ABS_n na sygnał wzorcowy AS_n . Jeśli nie, sygnał pozostaje bez zmian.

Na Fig. 2 można zauważyć, że podmianie uległy fragmenty sygnału z przesłuchem odpowiadające częstotliwościom 300 Hz i 400 Hz jako znacząco różniące się od wzorca. Fragment dla częstotliwości 100 Hz i 200 Hz pozostał bez zmian, ponieważ amplituda sygnału z przesłuchem mieściła się w zakresie dopuszczalnego przyjętego w tym przykładzie wykonania progu. W zależności od potrzeb, możliwe jest wykorzystanie różnych wartości progowych.

Pomimo wielu zdarzeń A wybiera się jedno z nich i oblicza się AS_n . Stanowi ono wzorzec dla porównań z pozostałymi zdarzeniami B a dokładniej porównania z wartościami ABS_n .

Niewątpliwą zaletą tego rozwiązania jest fakt, iż dzięki takiemu rozwiązaniu, każde z rejestrowanych sygnałów zachowuje swój unikatowy charakter. Dzieje się tak gdyż nie jest zmieniany cały sygnał AB (z przesłuchem) lecz jedynie składowa.

Zastrzeżenia patentowe

1. Sposób eliminacji przesłuchów przy realizacji nagrań dźwiękowych, w którym za pomocą pierwszego urządzenia rejestrującego sygnał rejestruje się sygnał z pierwszego źródła sygnału, jednocześnie w różnych odstępach czasu przez pierwsze urządzenie rejestrujące sygnał rejestruje się niechciane sygnały z co najmniej jednego innego źródła sygnału, **znamienny tym**, że z zarejestrowanego sygnału z pierwszego źródła sygnału wybiera się wzorcowe impulsy i dla wybranego wzorcowego impulsu wykonuje się wielopakietową transformatę falkową sygnału i otrzymuje się wzorcowe wartości kryterialne analizowanego fragmentu dla każdej ze skal, następnie wielopakietowej transformacie falkowej poddaje się wszystkie pozostałe sygnały z przesłuchem, a następnie porównuje się otrzymane po transformacie falkowej fragmenty sygnału z przesłuchem z odpowiednimi fragmentami otrzymanymi po transformacie falkowej sygnału wzorcowego, przy czym gdy fragmenty sygnału z przesłuchem znajduje się poza z góry określonym obszarem następuje zamiana fragmentu sygnału z przesłuchem na fragment sygnału bez przesłuchu.

2. Sposób według zastrz. 1, **znamienny tym**, że urządzeniem rejestrującym sygnał jest mikrofon.

3. Sposób według zastrz. 1, **znamienny tym**, że z góry określony obszar wynosi 10% wartości sygnału wzorcowego.

Rysunki

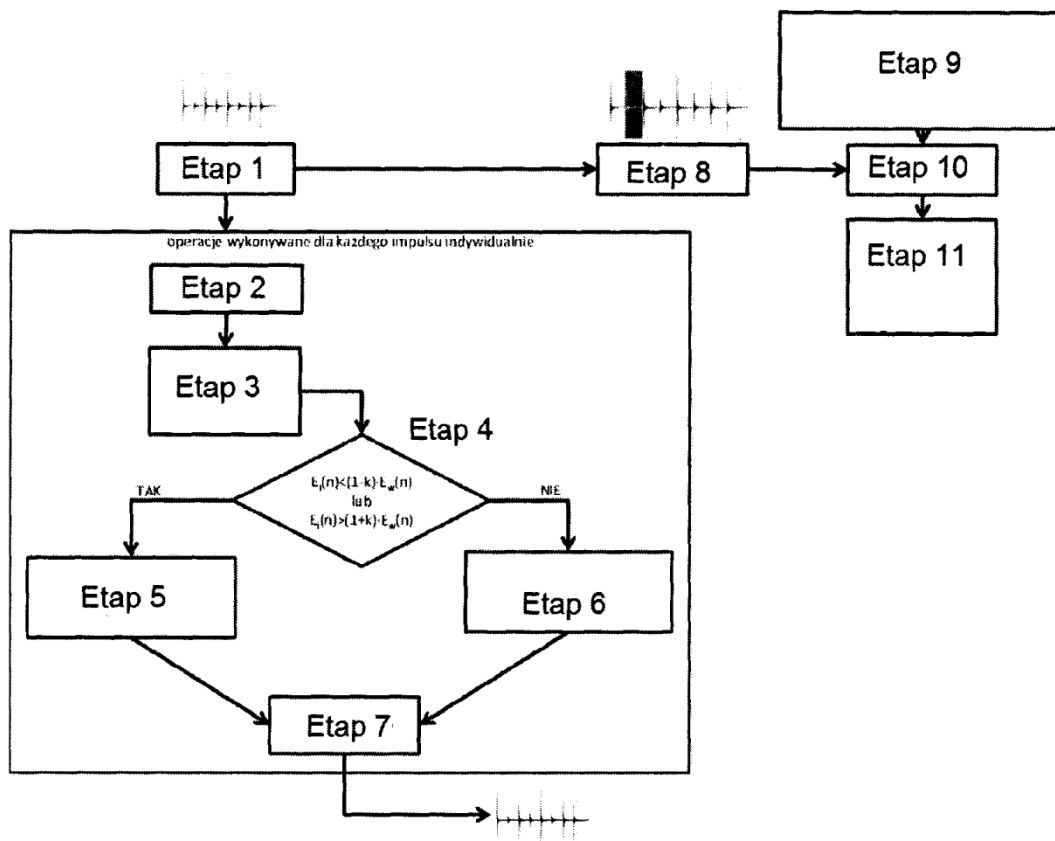


Fig. 1

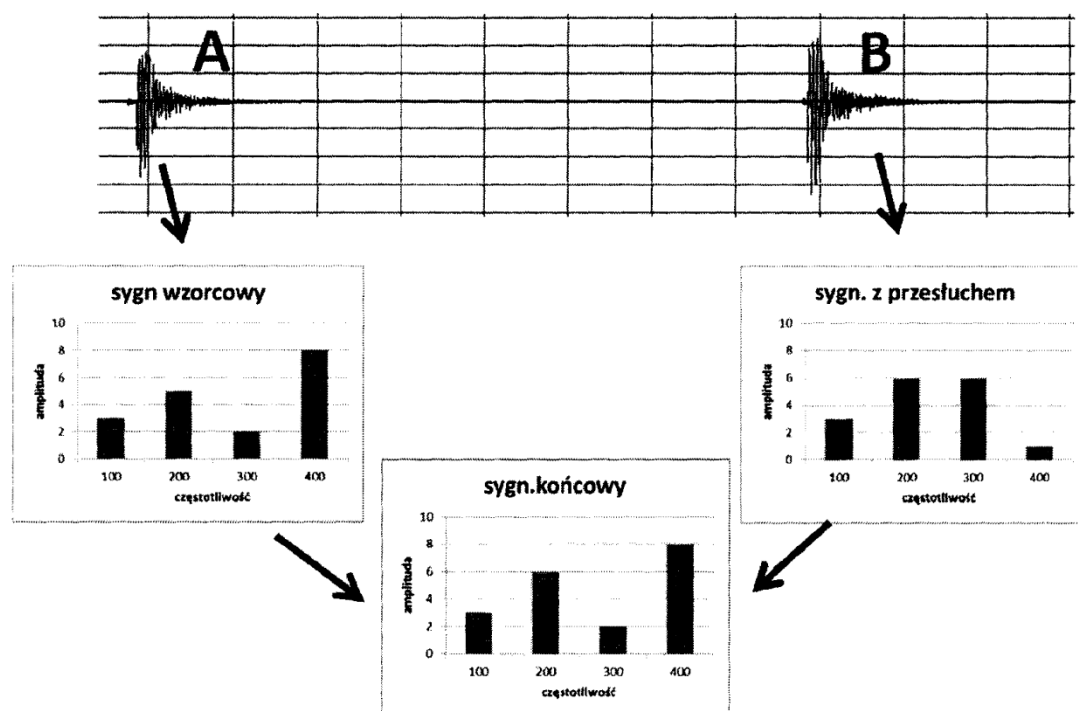


Fig. 2

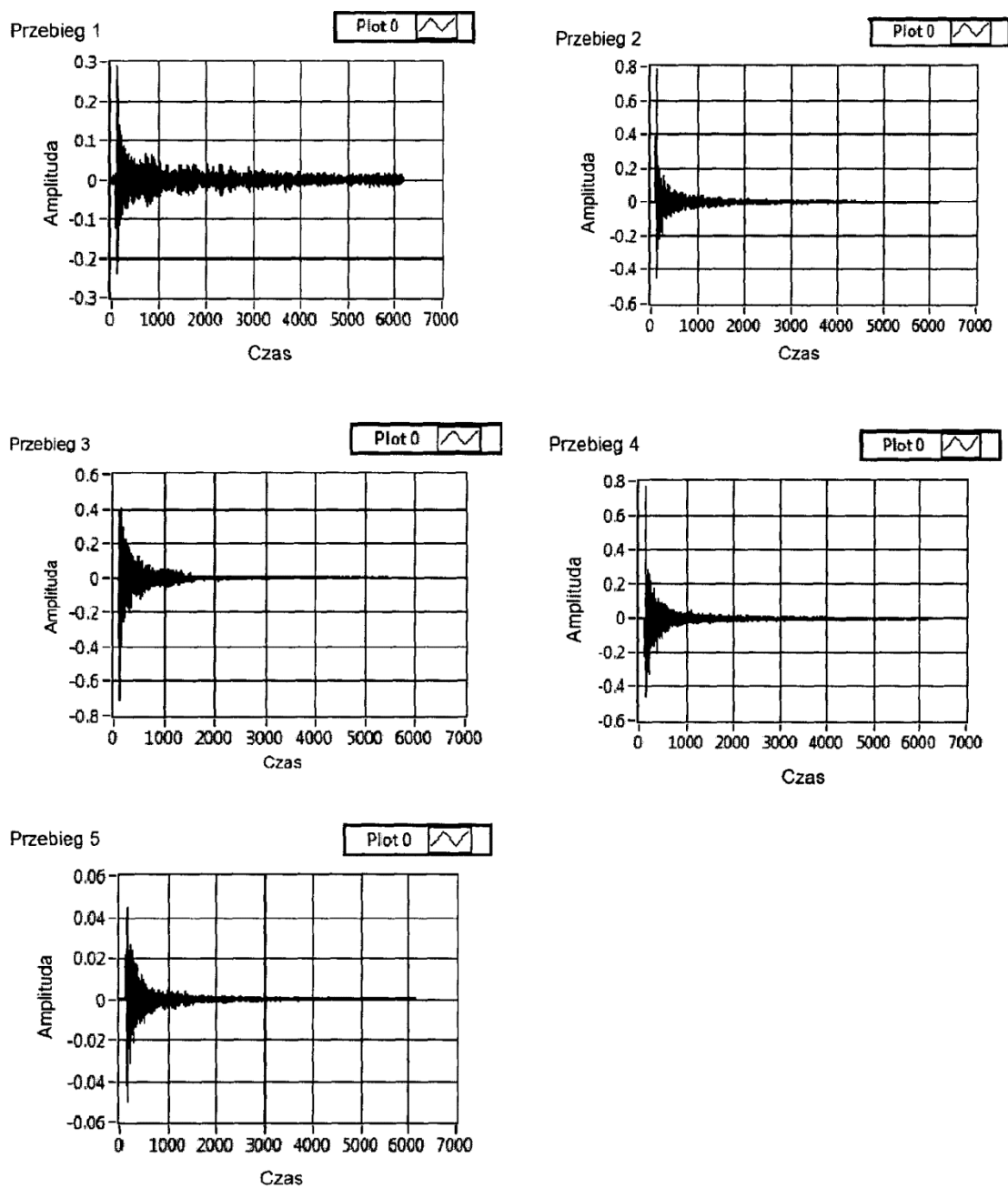


Fig. 3

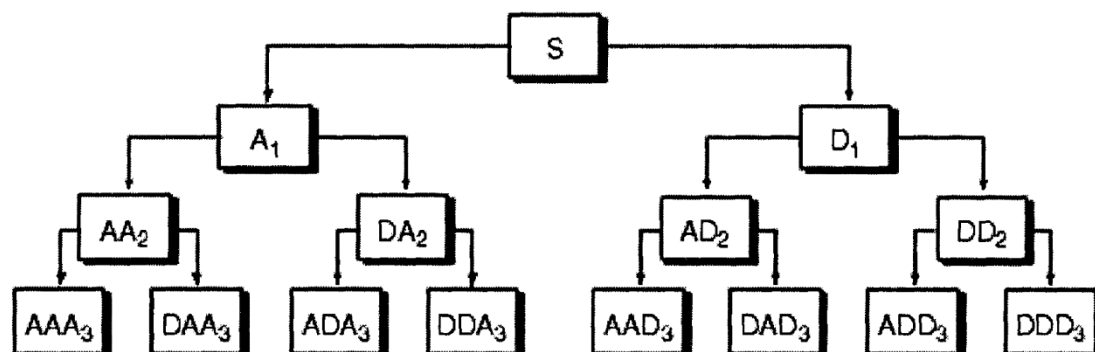


Fig. 4

