

Fumio Adam Okazaki*, Włodzimierz Kasprzak

Rozplatanie sygnałów mowy z ich mieszanin przy założeniu stałej przekątnej macierzy mieszającej

1 Wprowadzenie

Usługowe roboty muszą operować w środowiskach, w którym **poruszają się ludzie**, w tym też mogą **kooperować z ludźmi** przy realizacji swoich zadań. Dlatego też niezbędne jest wyposażenie robota w czujniki sygnału mowy - **mikrofony** - umieszczone zarówno na pojazdach jak i na stałe w środowisku, umożliwiające analizę sygnału mającą celu **detekcję sygnału mówcy** podającego komendy głosem.

Wychodzimy z założenia, że „nie ma lepszych danych od większej liczby danych”. Dlatego też metody pozyskiwania danych zakładają istnienie nadmiarowych czujników mowy. Zakładamy działanie robota w częściowo znanym środowisku wewnętrznym, w którym rozmieszczono wiele mikrofonów - na platformach mobilnych i w otoczeniu robotów. Zagadnienie rekonstrukcji właściwego, niezakłóconego sygnału mowy z pewnej ilości bodźców (zakłóconych sygnałów) modelowane jest jako problem **ślepego rozplatania źródeł** (BSD – *Blind Source Deconvolution*) [1-3], rozwiązywany przez algorytm **uczenia bez nauczyciela** macierzowej funkcji przetwarzającej **sygnały wektorowe**.

Opracowywana metoda „ślepego rozplatania” dźwięków może należeć do jednej z dwóch zasadniczych grup:

- wielokrotne wykorzystanie metody ślepej separacji źródeł w przestrzeni Fouriera [4, 5, 7];

Główne problemy wymagające rozwiązania w tej grupie metod to [5]: (1) zapewnienie identycznych permutacji sygnałów dla wszystkich cząstkowych problemów separacji, realizowanych niezależnie dla składowych w dziedzinie częstotliwości; oraz (2) zapewnienie właściwej relacji amplitud sygnałów w tych cząstkowych problemach separacji;

- uogólnienie metody ślepej separacji w dziedzinie czasu – zamiast skalarnych współczynników należy nauczyć współczynniki wektorowe – współczynniki filtrów FIR [2, 6, 8, 9];

Główne problemy wymagające rozwiązania w tej grupie metod to: (1) modelowanie niezerowych auto-korelacji sygnałów źródeł w dziedzinie czasu, (2) dobór funkcji nieliniowych oraz (3) dobór ilości kanałów i długości filtrów rozplatających.

W przeszłości autorzy zaproponowali oryginalne podejście do rozplatania mieszanin

* Instytut Automatyki i Informatyki Stosowanej Politechniki Warszawskiej, ul. Nowowiejska 15/19, 00-665 Warszawa. E-mail: A.Okazaki@elka.pw.edu.pl, W.Kasprzak@ia.pw.edu.pl

obrazów [10, 11], wykorzystując istnienie dominującej częstotliwości w spektrum takich sygnałów. Możliwe staje się uniknięcie (w rozsądnym przybliżeniu) problemu nieznanej permutacji i skali składowych spektrum, uzyskanych z zastosowania metody „separacji mieszanin punktowych” do każdej częstotliwości oddzielnie. Jednak sygnały mowy mają relatywnie nieprzewidywalne i zróżnicowane spektrum częstotliwościowe w porównaniu z obrazami naturalnymi. Podjęliśmy więc trud opracowania metody rozplatania sygnałów wywodzącej się z drugiego typu podejść. Da to nam narzędzie dla detekcji nie-zakłócanego sygnału mówcy - odtworzenia idealnego sygnału mowy z mieszaniny dźwięków i szumu docierających do mikrofonów.

W rozdziale drugim przedstawiono problem BSD i opis dwóch algorytmów proponowanych w niniejszej pracy (metody CPA i CDA dla problemu BSD). Następny rozdział najpierw omawia przykładowe wyniki eksperymentalne, jakie zostały uzyskane za pomocą przedstawionych w pracy algorytmów. Całość prac podsumowano w ostatnim rozdziale.

2 Algorytmy CPA i CDA dla rozplatania źródeł

W pracy omówiono i zaimplementowano dwa algorytmy wielo-kanalowego rozplatania źródeł (problem MBD) z ich mieszanin – tzw. algorytm CPA (*Constant Power separation Algorithm* – algorytm separacji stałoenerygetycznej) i CDA (*Constant Diagonal separation Algorithm* – algorytm separacji stałodiagonalnej), których ideę podano w pracy [8] i rozwinięto w [12]. Oba algorytmy (CPA i CDA) oparte są na założeniu tzw. stałodiagonalnej macierzy mieszającej.

2.1 Model algorytmów z wykorzystaniem stałodiagonalnej macierzy mieszającej

Zakładany model mieszania źródeł: sygnały źródłowe $\{s_i(k) \ (i=1, \dots, n)\}$ są statystycznie niezależne, a ich mieszaniny $\{x_i(k) \ (i=1, \dots, n)\}$ są (w dyskretnym i skończonym czasie) splotami źródeł $\{s_i(k)\}$ wyrażonymi poniższym wzorem [2]:

$$x_i(k) = \sum_{j=1}^n \sum_{l=0}^L h_{ij}(l) s_j(k-l) \quad (2.1)$$

gdzie L jest stopniem filtru FIR (liczbą opóźnień czasowych), a $\{h_{ij}\}$ współczynnikami macierzy mieszającej.

Celem algorytmu separacji jest znalezienie współczynników

$$\mathbf{w}_{ij}^T = [w_{ij}(0), \dots, w_{ij}(q)] \quad (2.2)$$

dla każdego j -tego wejścia i i -tego wyjścia.

Po zebraniu wektorów współczynników w macierze otrzymamy macierzowe postacie modeli mieszania i separacji:

$$\begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix} = \begin{pmatrix} h_{11}^T & \cdots & h_{1n}^T \\ \vdots & \ddots & \vdots \\ h_{n1}^T & \cdots & h_{nn}^T \end{pmatrix} \begin{pmatrix} s_1(k) \\ \vdots \\ s_n(k) \end{pmatrix} \quad (2.3(a))$$

$$\begin{pmatrix} y_1(k) \\ \vdots \\ y_n(k) \end{pmatrix} = \begin{pmatrix} w_{11}^T & \cdots & w_{1n}^T \\ \vdots & \ddots & \vdots \\ w_{n1}^T & \cdots & w_{nm}^T \end{pmatrix} \begin{pmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{pmatrix} \quad (2.3(b))$$

Należy zauważyć, że w powyższych modelach wszystkie elementy **po prawej** stronie równań są wektorami danych dla opóźnień czasowych. M.in.:

$$s_i^T(k) = [s_i(k), \dots, s_i(k-p)] \quad (2.4(a))$$

$$x_i^T(k) = [x_i(k), \dots, x_i(k-q)] \quad (2.4(b))$$

Po przejściu do dziedzin transformaty Z otrzymamy:

$$F(z) = \sum_{k=0}^K f(k) z^{-k}$$

$$X_i(z) = \sum_{j=1}^n H_{ij}(z) S_j(z); \quad \mathbf{X}(z) = \mathbf{H}(z) \mathbf{S}(z); \quad (2.5)$$

$$\text{gdzie: } \mathbf{H}(z) = \begin{pmatrix} H_{11}(z) & \cdots & H_{1n}(z) \\ \vdots & \ddots & \vdots \\ H_{n1}(z) & \cdots & H_{nn}(z) \end{pmatrix}, \quad \mathbf{x}(z) = \begin{pmatrix} x_1(z) \\ \vdots \\ x_n(z) \end{pmatrix}, \quad \mathbf{s}(z) = \begin{pmatrix} s_1(z) \\ \vdots \\ s_n(z) \end{pmatrix} \quad (2.6)$$

Problem rozplatania możemy uprościć zakładając, że macierz mieszająca jest w jakiś sposób znormalizowana. Rozpatrywana tu stało-diagonalna macierz mieszająca przyjmuje następującą postać [8]:

$$\mathbf{H}(z) = \mathbf{H}_a(z) \mathbf{H}_b(z) \quad (2.7)$$

$$\mathbf{H}_a(z) = \begin{pmatrix} 1 & \frac{H_{21}(z)}{H_{11}(z)} & \cdots & \frac{H_{1n}(z)}{H_{11}(z)} \\ \frac{H_{21}(z)}{H_{11}(z)} & \ddots & & \vdots \\ \vdots & & \ddots & \frac{H_{(n-1)n}(z)}{H_{11}(z)} \\ \frac{H_{n1}(z)}{H_{11}(z)} & \cdots & \frac{H_{n(n-1)}(z)}{H_{(n-1)(n-1)}(z)} & 1 \end{pmatrix}, \quad \mathbf{H}_b(z) = \begin{pmatrix} H_{11}(z) & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & H_{nn}(z) \end{pmatrix} \quad (2.8)$$

stąd:

$$\mathbf{X}(z) = \mathbf{H}(z) \mathbf{S}(z) = \mathbf{H}_a(z) \mathbf{H}_b(z) \mathbf{S}(z) = \mathbf{H}_a(z) \mathbf{S}'(z) \quad (2.9)$$

gdzie:

$$\mathbf{S}'(z) = \begin{pmatrix} s_1'(z) \\ \vdots \\ s_n'(z) \end{pmatrix} = \begin{pmatrix} H_{11}(z) s_1(z) \\ \vdots \\ H_{nn}(z) s_n(z) \end{pmatrix} \quad (2.10)$$

To założenie oznacza, że każde źródło dociera do dokładnie jednego sensora w postaci nie-filtrowanej. Częściowym uzasadnieniem tego założenia jest fakt, że proces separacji może odzyskać źródło jedynie z dokładnością do skali. Pominięcie opóźnień czasowych w procesie mieszania jednego źródła dla każdego sensora jest jednak istotnym uproszczeniem generalnego problemu BSD. To założenie może być usprawiedliwione w praktyce bliskim położeniem sensora względem dokładnie jednego źródła.

W procesie rozplatania poszukiwana będzie macierz separująca, reprezentująca filtr FIR:

$$\mathbf{W}(z) = \mathbf{P}(z) \text{cof}(\mathbf{H}(z)) \quad (2.11(a))$$

pod warunkiem, że macierz mieszająca ma niezerowy wyznacznik ($\det(\mathbf{H}(z)) \neq 0$).

Współczynnik $\text{cof}(\mathbf{H}(z))$ związany jest z macierzą odwrotną $\mathbf{H}^{-1}(z)$ zależnością:

$$\text{cof}(\mathbf{H}(z)) = \det(\mathbf{H}(z))\mathbf{H}^{-1}(z) \quad (2.11(b))$$

natomiast $\mathbf{P}(z)$ jest macierzą permutacji wierszy. Stąd:

$$\begin{aligned} \mathbf{Y}(z) &= \mathbf{W}(z)\mathbf{X}(z) = \mathbf{W}(z)\mathbf{H}(z)\mathbf{S}(z) = \mathbf{P}(z)\text{cof}(\mathbf{H}(z))\mathbf{H}(z)\mathbf{S}(z) = \\ &= \mathbf{P}(z)\det(\mathbf{H}(z))\mathbf{H}^{-1}(z)\mathbf{H}(z)\mathbf{S}(z) = \mathbf{P}(z)\det(\mathbf{H}(z))\mathbf{S}(z) \end{aligned} \quad (2.12)$$

Ponieważ celem procesu separacji-rozplatania jest statystyczna niezależność sygnałów wyjściowych, dla każdego z opóźnień czasowych $l=(l_1, \dots, l_2)$ (np. $l_1=0, l_2=L$), przyjmowana jest następująca funkcja kosztu (celu):

$$C = \sum_{i=1}^n \sum_{j=1, j \neq i}^n \sum_{l=l_1}^{l_2} r_{y_i y_j}^2(l) \quad (2.13)$$

gdzie $r_{y_i y_j}^2(l) = E^2 \{ f[y_i(t)] g[y_j(t-l)] \}$ dla sygnałów o zerowej wartości średniej, a $f[y]$ i $g[y]$ są odpowiednio dobranymi parami funkcji:

$$\begin{aligned} f[y] &= y^3 ; g[y] = y \text{ (dla tzw. pod-gaussowskich sygnałów) lub} \\ f[y] &= y ; g[y] = \tanh(y) \text{ (dla tzw. nad-gaussowskich sygnałów).} \end{aligned}$$

2.2 Dwa algorytmy BSD dla stało-diagonalnej macierzy mieszającej

Z warunku minimalnej funkcji celu mogą być wyprowadzone iteracyjne reguły modyfikacji wag (elementów macierzy \mathbf{W}) [12]. Przedstawimy dwa podejścia zapobiegające zapadnięciu się macierzy wag $W(z)$ do 0 (tzn. chroniące algorytm przed oczywistym rozwiązaniem dla wartości macierzy wag równych 0): podejście **stałodiagonalne** i **stałoenergetyczne**.

Algorytm Stałodiagonalny (CDA) zapobiega zapadnięciu się macierzy wag do zera poprzez z góry założone ustawienie wag $\mathbf{w}_{ii}^T = [1, 0, \dots, 0]$ dla $i=1 \dots n$.

Algorytm CDA BSD

(1) Inicjalizuj wagi \mathbf{w}_i

(2) **REPEAT**

for $l=1 \dots n$

for $m=1 \dots n$

{ jeśli $l \neq m$, aktualizuj \mathbf{w}_{lm} wg wzoru (2.14) }

UNTIL wagi się ustabilizują.

Aktualizacja wagi \mathbf{w}_{lm} następuje według reguły:

$$\mathbf{w}_{lm} = \left(\sum_{j \neq l} \sum_{c=1}^n \sum_{d=1}^n \mathbf{A}_{jc}^T \mathbf{R}_{x_m x_c} \mathbf{R}_{x_m x_d}^T \mathbf{A}_{jd} \right)^{-1} \cdot \left(\sum_{j \neq l} \sum_{c=1}^n \sum_{b \neq m} \sum_{d=1}^n \mathbf{A}_{jc}^T \mathbf{R}_{x_m x_c} \mathbf{R}_{x_b x_d}^T \mathbf{A}_{jd} \mathbf{w}_{lb} \right)^{-1} \quad (2.14)$$

gdzie \mathbf{A} jest macierzową funkcją wag \mathbf{w} , natomiast \mathbf{R} jest macierzą funkcji wzajemnej niezależności sygnałów zmierzonych x . A mianowicie \mathbf{A}_{jc} oznacza następującą macierz złożoną z tych elementów szukanych $\{w_{ij}(q)\}$, które wykorzystywane są podczas obliczeń dla l, m, j :

$$\mathbf{A}_{jc} = \begin{pmatrix} w_{jc}(q) & 0 & \dots & 0 & 0 \\ w_{jc}(q) & w_{jc}(q) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ w_{jc}(1) & w_{jc}(2) & \dots & w_{jc}(q) & 0 \\ w_{jc}(0) & w_{jc}(1) & \dots & w_{jc}(q-1) & w_{jc}(q) \\ 0 & w_{jc}(0) & \dots & w_{jc}(q-2) & w_{jc}(q-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & w_{jc}(0) & w_{jc}(1) \\ 0 & 0 & \dots & 0 & w_{jc}(0) \end{pmatrix} \quad (2.15)$$

$\mathbf{R}_{x_a x_c}$ oznacza następującą macierz korelacji par sygnałów przy znanym opóźnieniu wzajemnym l :

$$r_{x_i x_j}(l) = E\{f[x_i(k)]g[x_j(k+l)]\} \quad (2.16(a))$$

a funkcje nieliniowe $f(x)$ i $g(x)$ definiowane są podobnie jak w równaniu (2.13).

$$r_{x_i x_j}^{(l)} = [r_{x_i x_j}(l-q), \dots, r_{x_i x_j}(l+q)]^T \quad (2.16(b))$$

$$R_{x_a x_c}(l) = \begin{pmatrix} r_{x_a x_c}^{(l)} & \dots & r_{x_a x_c}^{(l-q)} \\ \vdots & \ddots & \vdots \\ r_{x_a x_c}^{(l+q)} & \dots & r_{x_a x_c}^{(l)} \end{pmatrix} \quad (2.16(c))$$

Funkcja kosztu dla algorytmu CDA przyjmuje następującą postać:

$$C = \left(\sum_{i=1}^n \sum_{j=1, j \neq i}^n \sum_{a,b,c,d=1}^n w_{ia}^T A_{jc}^T R_{x_a x_c} R_{x_b x_d}^T A_{jd} w_{ib} \right) \quad (2.17(a))$$

Minimalizacja wartości C względem współczynników w_{lm} jest celem algorytmu separacji:

$$\frac{\partial C}{\partial w_{lm}} = 0. \quad (2.17(b))$$

Algorytm Stałoenergetyczny (CPA BSD) zapobiega zapadnięciu się macierzy wag poprzez ustawienie współczynników autokorelacji sygnałów dla zerowych opóźnień na pewną ustaloną wartość K :

$$R_{y_i y_i}[0] = K, \quad (\text{dla } i=1 \dots n) \quad (2.18)$$

Algorytm CPA BSD

(1) Inicjalizuj wagi w_i

(2) **REPEAT**

for $i=1 \dots n$

{ - skonstruuj G_i ; według wzoru (2.19)

- zamień w_i z przeskalowanym wektorem własnym $R_x^{-1}(0) G_i$ z najmniejszą wartością własną dla $w_i^T R_x(0) w_i = K$;

UNTIL: wagi się ustabilizują

W powyższym algorytmie oznaczono

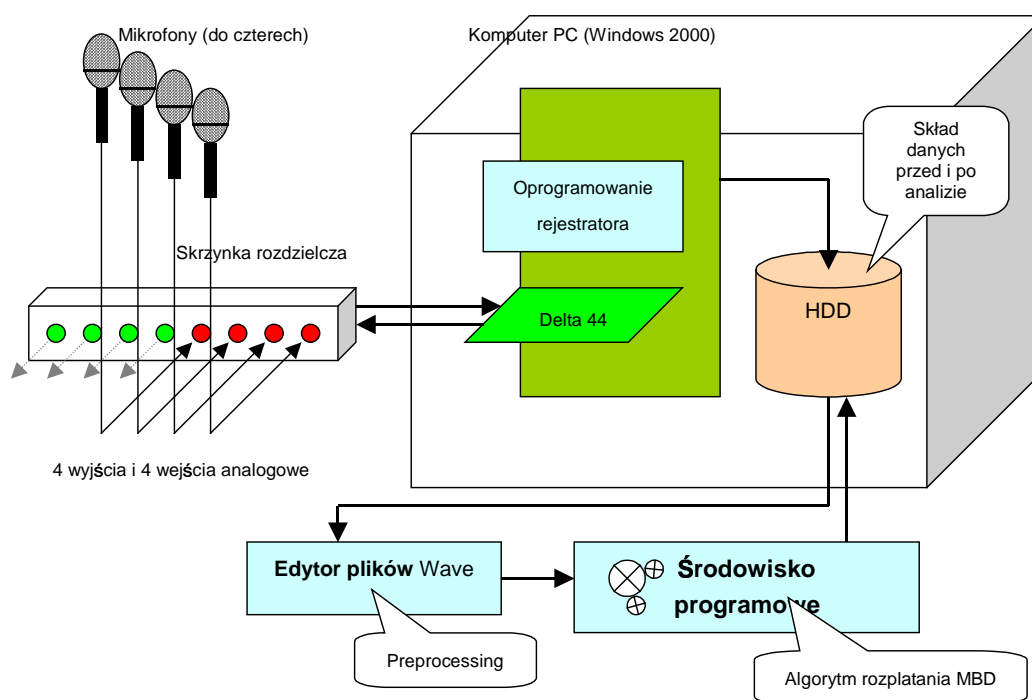
$$G_i = \sum_{j=1, j \neq i}^n \tilde{A}_j^T M_x M_x^T \tilde{A}_j \quad (2.19)$$

gdzie \mathbf{A} jest macierzą funkcją wag w_j , natomiast \mathbf{M} jest macierzą funkcji wzajemnej niezależności zmierzonych sygnałów wejściowych x .

3 Wyniki eksperymentalne

Badania eksperymentalne przeprowadzono na następującym stanowisku [12] (Rys. 1). Jako wielokanałowa karta dźwiękowa wybrana została karta Delta 44 firmy *M-Audio* [13]. Karta została zainstalowana na złączu PCI komputera osobistego z procesorem Intel Pentium 2.4GHz z pamięcią podręczną 512kB i szyną FSB taktowaną z częstotliwością 533MHz. Dołączono 4 mikrofony dynamiczne C608 produkcji amerykańskiej firmy *Shure*, pracujące w przedziale częstotliwości 50 - 15 kHz [14].

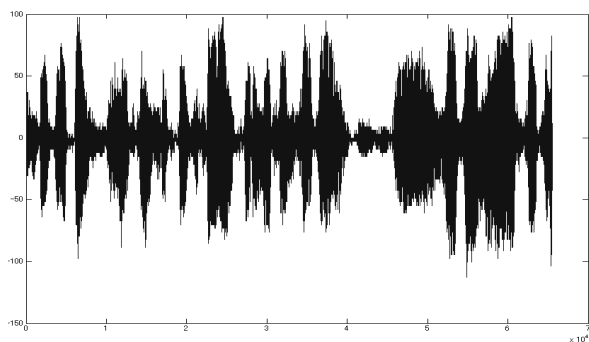
Parametry mikrofonów pozwalają dobrze odbierać sygnały mowy mówionej z niewielkich odległości, tj. do 30 centymetrów. To pozwala uznać teoretyczne założenia algorytmów stało-diagonalnych za spełnione w praktyce.



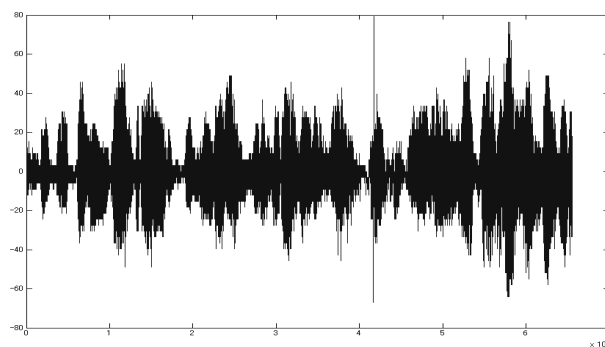
Rysunek 1. Schemat ideowy stanowiska

Stacja robocza została wyposażona w oprogramowanie do jednoczesnej rejestracji dźwięku z wielu kart dźwiękowych i/lub wielu kanałów *n-Track Studio*, edycja 24 bitowa w wersji 3.3 [15], a także w prosty edytor plików Wave o nazwie *Audacity* [16].

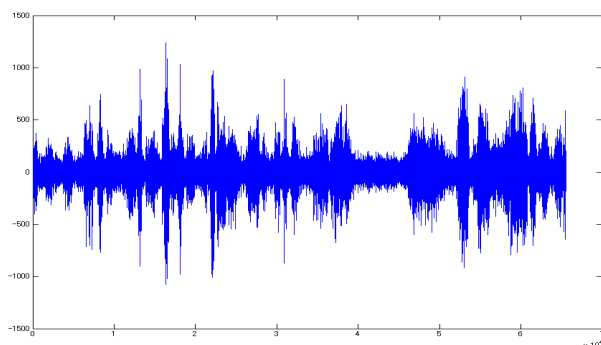
Próbka dźwiękowa jest nagrana z jakością 16 bit / 44 kHz, jednak zostaje później cyfrowo przepróbkowana do częstotliwości 12 kHz, ze względów czysto praktycznych i obliczeniowych (czas trwania analizowanej próbki to ok. 5 i pół sekundy).



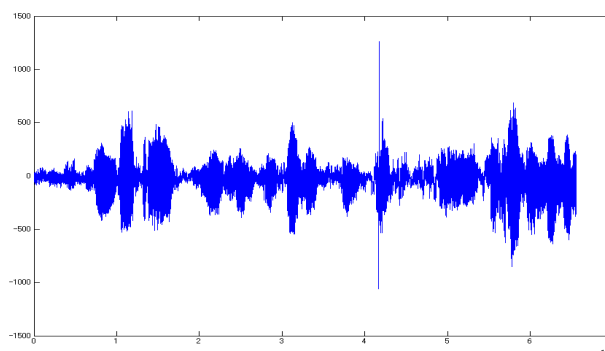
Rysunek 2. Pierwszy sygnał nagrany z mikrofonu (mieszanina x_1).



Rysunek 3. Drugi sygnał nagrany z mikrofonu (mieszanina x_2).



Rysunek 4. Pierwszy sygnał odseparowany metodą CDA (estymacja y_1).



Rysunek 5. Drugi sygnał odseparowany metodą CDA (estymacja y_2).

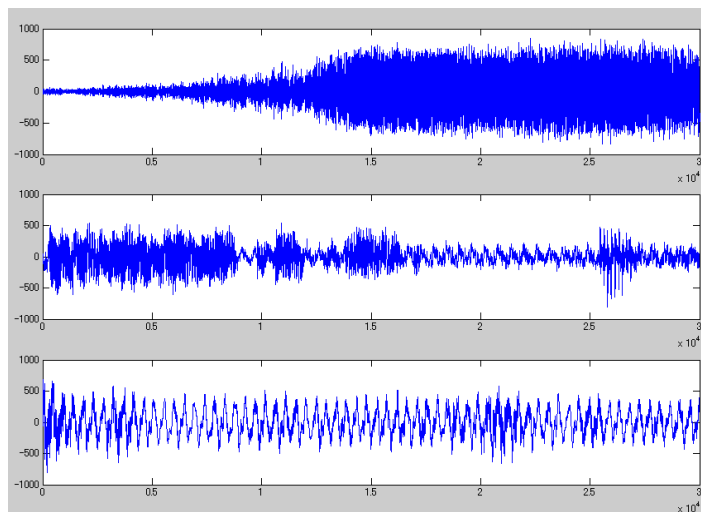
W ramach testów stanowiska oraz zaimplementowanego algorytmu, przeprowadzono również eksperymenty ze sztucznie utworzonymi mieszaninami sygnałów dźwiękowych. Poniżej znajdują się kolejno sygnały źródłowe, ich mieszaniny i odseparowane sygnały, uzyskane w wyniku dwóch procesów mieszania i separacji próbnych sygnałów dźwiękowych. Pierwsza separacja została przeprowadzona algorytmem stałoenergetycznym CPA (3 sygnały), natomiast druga rozszerzonym algorytmem stałodiagonalnym ECDA (4 sygnały).

W obu przypadkach wykorzystano naturalne sygnały dźwiękowe, zakodowane z częstotliwością 12000 Hz i rozdzielczością 16 bitów na próbkę. Liczba próbek wynosiła 30000, czyli każdy sygnał był długości dokładnie 2.5 sekundy.

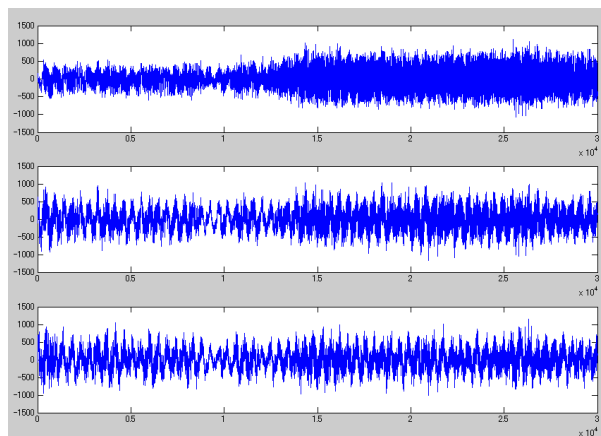
Obliczenia z wykorzystaniem algorytmu CPA do separacji 3 sygnałów źródłowych trwały zaskakująco krótko, ok. 2 sekund (liczba iteracji: 94). Mimo krótkiego czasu obliczeń, jakość separacji jest bardzo dobra, w zupełności zadowalająca.

Oczywiście, algorytm odtworzył sygnały źródłowe z dokładnością do ich kolejności i znaku (łatwo zauważyć, że np. drugi sygnał źródłowy został odseparowany jako pierwszy z ujemnym znakiem), co jest naturalnie nieuniknione (między innymi dlatego, że już podczas samego procesu mieszania informacja o kolejności i znakach sygnałów jest tracona).

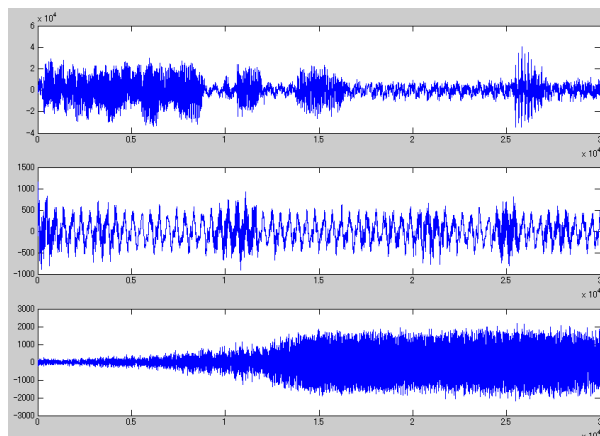
Separacja algorytmem ECDA 4 mieszanin trwała już znacznie dłużej (ok. 20 sekund), natomiast licznik iteracji wskazał wartość 3019. Warto jednak podkreślić wysoką jakość separacji w tym przypadku.



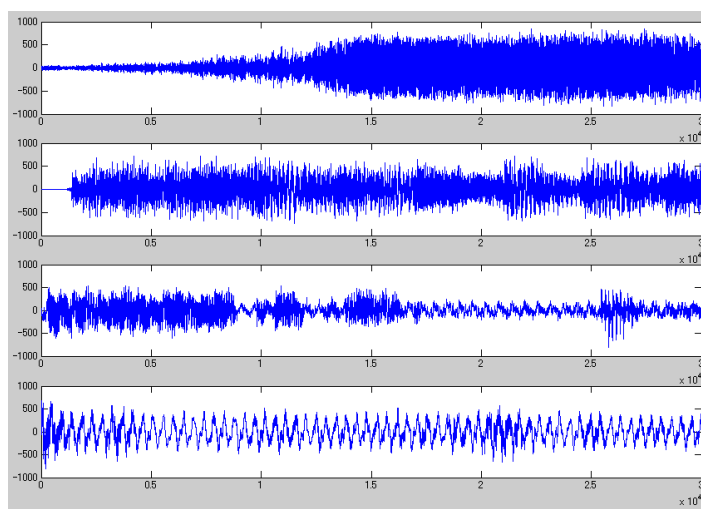
Rysunek 6. 3 sygnały źródłowe, wykorzystane w algorytmie CPA.



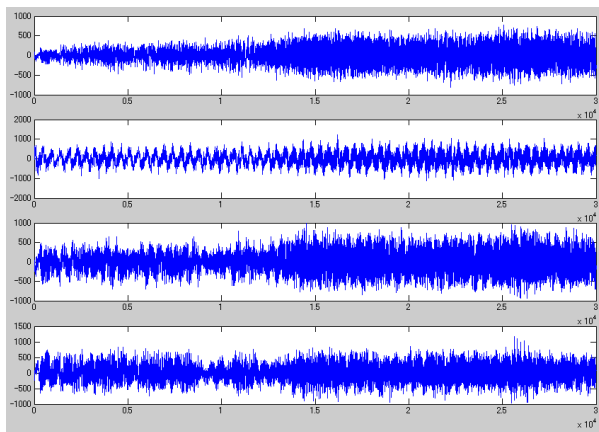
Rysunek 7. 3 mieszaniny, podane jako wejście do algorytmu CPA (długość filtra mieszającego: 2)



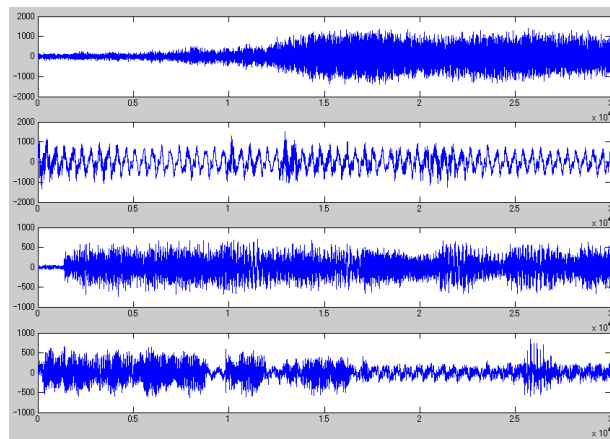
Rysunek 8. 3 sygnały odseparowane algorytmem CPA



Rysunek 9. 3 sygnały źródłowe, wykorzystane w algorytmie ECDA



Rysunek 10. 3 mieszanki, podane jako wejście do algorytmu ECDA (długość filtra mieszającego: 3)



Rysunek 11. 3 sygnały odseparowane algorytmem ECDA.

4 Podsumowanie

Opracowano metodę **analizy wielu sygnałów**, które pochodzą z **czujników dźwięku** (zawierających ludzką **mowę**), występujących w **robotów** usługowych. W szczególności przedmiotem badań była **detekcja** niezakłóconego **sygnału mowy** w warunkach swobodnego wydawania komend przez człowieka (mikrofony rozmieszczone w pomieszczeniu a nie związane z mówcą). Dla rozwiązania problemu zastosowano rekonstrukcję użytecznego sygnału z mieszanin wielu źródeł metodą rozplatania sygnałów wektorowych.

Zaproponowano dwa algorytmy rozplatania mieszanin mowy. Przebadano praktycznie oba algorytmy uzyskując dobre wyniki separacji sygnałów mowy. W praktyce poprawę wyników można osiągnąć dzięki optymalnemu doborowi parametrów separacji (optymalizowanych uprzednio dla mieszanin sztucznych). Nie bez wpływu były też między innymi parametry mikrofonów użytych na stanowisku, wymuszające silne wzmocnienie zarejestrowanego sygnału na etapie przetwarzania wstępnego, w wyniku czego traci się pewną istotną część poziomów kwantyzacji, a tym samym część informacji zawartej w sygnałach (co widać również na rysunkach 2 i 3). Warto nadmienić, iż wspomniany wcześniej proces prze-próbkowania sygnałów wejściowych z częstotliwości 44 kHz do 12 kHz nie wprowadził do sygnału znaczących zniekształceń. W ramach dalszych prac przewiduje się testy na większej ilości kanałów.

Podziękowanie

Niniejszą pracę wykonano w ramach projektu badawczego finansowanego przez grant MNiI - 4T11A 003 25.

Literatura

- [1] A. Hyvarinen, J. Karhunen, E. Oja: *Independent Component Analysis*, John Wiley & Sons, New York etc., 2001.
- [2] A. Cichocki, S. Amari: *Adaptive Blind Signal and Image Processing*, John Wiley, Chichester, UK, 2002.
- [3] S. Amari, A. Cichocki, S. Makino, N. Murata. (eds.): *Fourth Int. Symposium on Independent Component*

Analysis and Blind Signal Separation - ICA'2003. Proceedings. Nara, Japan, 2003, NTT Comm. Science Labs. ISBN 4-9901531-0-3.

- [4] P. Smaragdis: Blind separation of convolved mixtures in frequency domain. *Neurocomputing*, vol. 22 (1998), 21-34.
- [5] N. Araki et al.: Fundamental limitation of frequency domain blind source separation for convolved mixture of speech. *Proceedings ICASSP2001*, vol. 5, May 2001, 2737-2740.
- [6] N. Murata, S. Ikeda, A. Ziehe: An approach to blind source separation based on temporal structure of speech signal. *Neurocomputing*, vol. 41(2001), No. 4, 1-24.
- [7] H. Saruwatari, S. Kurita, K. Takeda: Blind source separation combining frequency-domain ICA and beamforming. *Proceedings ICASSP2001*, May 2001, 2733-2736.
- [8] D.C.B. Chan: *Blind Signal Separation*. Praca doktorska, University of Cambridge, Wydział Inżynieryjny, Cambridge, UK, January 1997.
- [9] W. Kasprzak, A. Cichocki, S. Amari: Blind Source Separation with Convolutional Noise Cancellation, *Neural Computing and Applications*, Springer-Verlag London Ltd., vol. 6(1997), 127-141.
- [10] W. Kasprzak, A. Okazaki: *Blind deconvolution of timely-correlated sources by homomorphic filtering in Fourier space*, Fourth Int. Symposium on Independent Component Analysis and Blind Signal Separation - ICA'2003, Nara, Japan, 2003, NTT Comm. Science Lab., 1029 – 1034.
- [11] W. Kasprzak: *Blind Source Deconvolution by Homomorphic Filtering in Fourier Space*. K. Wojciechowski (Ed.), ICCVG 2002 Conference Proceedings. International Conference on Computer Vision and Graphics, (25-29.09.2002, Zakopane), Silesian Univ. of Technology, Gliwice, ISBN 83-9176830-9, str. 431 - 436
- [12] A. Okazaki, W. Kasprzak: Ślepe Rozplatanie Sygnałów Mowy z ich mieszanin przy założeniu stałej przekątnej macierzy mieszającej, *Raport badawczy IAiS nr 04-02*, Instytut Automatyki i Informatyki Stosowanej Politechniki Warszawskiej, Warszawa, styczeń 2004.
- [13] M-Audio, Inc.: M-Audio Delta Series 44 User's Manual, USA 2000, <http://www.m-audio.com>
- [14] Shure Corporation: Shure C608 User's Manual, USA 2003.
- [15] F. Antonioli: Strona internetowa firmy *FASoft*, w tym program *n-Track Studio*: <http://www.fasoft.com/>
- [16] M. Burbeck, J. Haberman, D. Mazzoni: Strona internetowa projektu *Audacity*: <http://audacity.sourceforge.net/>