

Adam Fumio Okazaki*, Włodzimierz Kasprzak

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

1 Wprowadzenie

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], [4], [5], rozwiązywany przez algorytm uczenia bez nauczyciela macierzowej funkcji przetwarzającej sygnały wektorowe.

2 Algorytmy CPA i CDA dla rozplatania źródeł

W pracy omówiono implementację dwóch algorytmów wielo-kanałowego rozplatania źródeł (problem MBD) z ich mieszanin - tzw. algorytm CPA (Constant Power separation Algorithm - algorytm separacji stałoenergetycznej) i CDA (Constant Diagonal separation Algorithm - algorytm separacji stałodiagonalnej), których ideę podano w pracy [3] i rozwinięto w [10]. Oba algorytmy (CPA i CDA) oparte są na założeniu tzw. stałodiagonalnej macierzy mieszającej.

2.1 Problem MBD

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 [5]:

$$(2.1) \tag{1}$$

gdzie L jest stopniem filtru FIR (liczbą opóźnień czasowych), a h_{ij} współczynnikiem macierzy mieszającej. Celem algorytmu separacji jest odzyskanie nieznanych źródeł. Może to być zrealizowane przez kolejny splot (realizujące "semantycznie" - rozplecenie mieszanin), tym razem z macierzą elementów wektorowych (reprezentujących opóźnienia) W . Elementy tej, również nieznannej, macierzy to:

$$(2.2) \tag{2}$$

dla każdego j -tego wejścia (mieszaniny) i i -tego wyjścia (estymowanego źródła). Po zebraniu wektorów współczynników w macierze otrzymamy macierzowe postacie modeli mieszania i separacji:

$$(2.3) \tag{3}$$

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

*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

2.2 Założenie stałodiagonalnej macierzy mieszającej

Problem rozplatania możemy uprościć zakładając, że macierz mieszająca jest w jakiś sposób znormalizowana. Rozpatrywana tu stało-diagonalna macierz mieszająca przyjmuje następującą postać [3]:

$$(2.3) \tag{4}$$

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.

2.3 Algorytm stałodiagonalny i stałoenergetyczny

Z warunku minimalnej funkcji celu mogą być wyprowadzone iteracyjne reguły modyfikacji wag (elementów macierzy W) [10]. 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 $w_i T = [1, 0, \dots, 0]$ dla $i = 1 \dots n$.

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

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

3 Wyniki eksperymentalne

Badania eksperymentalne przeprowadzono na następującym stanowisku z wielokanałową kartą dźwiękową Delta 44 firmy M-Audio [13] z 4 mikrofonami dynamicznymi C608 firmy Shure, pracujące w przedziale częstotliwości 50 - 15 kHz [14].

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.

References

- [1] Amari S., Cichocki A., Makino S., Murata N. (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.
- [2] Araki et al.. Fundamental limitation of frequency domain blind source separation for convolved mixture of speech. *Proceedings ICASSP2001*, vol. 5, May 2001, 2737-2740.
- [3] D. C. B. Chan. *Blind Signal Separation*. Praca doktorska, University of Cambridge, Wydział Inżynieryjny, Cambridge, UK, January 1997.
- [4] A. Hyvarinen, J. Karhunen, E. Oja. *Independent Component Analysis*, John Wiley & Sons, New York etc., 2001.

- [5] A.Cichocki, S.Amari. *Adaptive Blind Signal and Image Processing*, John Wiley, Chichester, UK, 2002.
- [6] 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.
- [7] 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.
- [8] 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*, (September 25-29, 2002, Zakopane, Poland), str. 431 - 436. Silesian Univ. of Technology, Gliwice, ISBN 83-9176830-9.
- [9] N. Murata, S. Ikeda, A. Ziehe. An approach to blind source separation based on temporal structure of speech signal. *em Neurocomputing*, vol. 41(2001), No. 4, 1-24.
- [10] A.Okazaki, W.Kasprzak. Ślepe Rozplatanie Sygnałów Mowy z ich mieszanin przy założeniu stałej przekątnej macierzy mieszającej, *Raport badawczy IAiIS nr 04-02*, Instytut Automatyki i Informatyki Stosowanej Politechniki Warszawskiej, Warszawa, styczeń 2004.
- [11] H. Saruwatari, S. Kurita, K. Takeda. Blind source separation combining frequency-domain ICA and beamforming. *Proceedings ICASSP2001*, May 2001, 2733-2736.
- [12] P. Smaragdakis. Blind separation of convolved mixtures in frequency domain. *Neurocomputing*, vol. 22 (1998), 21-34.
- [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