Proceedings of the 6th WSEAS International Conference on Multimedia Systems & Signal Processing, Hangzhou, China, April 16-18, 2006 (pp146-150)

Interference Suppression in Diversity Wireless Receivers via ICA with Individual Adaptations

THOMAS YANG
Department of Electrical and Systems Engineering
Embry-Riddle Aeronautical University
600 S. Clyde Morris Blvd.
Daytona Beach, Florida 32114, USA

WASFY MIKHAEL
Department of Electrical and Computer Engineering
University of Central Florida
Orlando, Florida 32816, USA

Abstract: - In this paper, a novel Independent Component Analysis (ICA) algorithm is proposed with application to interference suppression in diversity receivers. For mobile and wireless communications, the new technique is an efficient baseband digital interference rejection method that results in simplified analog front-end. Simulations are performed for BPSK receivers assuming time-varying conditions. The results confirm the effectiveness of the presented technique. Also, the new technique is applicable in the presence of thermal noise.

Key-Words: - Interference suppression; ICA; Individual adaptation; Time varying channels; Diversity receiver

1 Introduction

For wireless receivers, to increase receiver flexibility and integration, most radio functionalities are required to be performed in the digital domain [1].

Interference suppression is an important task for communication systems. With the advancement of ADC technology, baseband digital interference rejection [2-6] is becoming increasingly feasible. This is an important step towards the realization of the future Software Defined Radio (SDR).

In this paper, mobile and wireless communications are considered in which the fading channels are often time-varying. A new ICA algorithm is proposed to perform interference suppression under these dynamic conditions. Simulation results for BPSK receivers indicate that the proposed technique is successful in achieving acceptable performance.

2 Receiver Structure and Signal Model

Fig. 1 shows the diversity BPSK receiver structure. Two antennas are used which, after identical downconversion stage, generate two baseband observations. Assume a co-channel interferer is present in the received signals.

The channel's fading coefficients are defined as:

$$f_{sk} = \alpha_{sk} e^{j\psi_{sk}} \quad (1)$$

$$f_{ik} = \alpha_{ik} e^{j\psi_{ik}} \quad (2)$$

where k=1, 2 is the antenna index; f_{sk}, f_{ik} are fading coefficients for the desired and interfering signals, respectively; α_{sk}, α_{ik} and ψ_{sk}, ψ_{ik} are the channel's amplitude and phase responses, respectively.

Let $s(t)$ and $i(t)$ denote the desired and interfering signals, respectively. Thus, the received signal of the kth antenna $r_k(t)$ can be expressed as:

$$r_k(t) = 2\mathrm{Re}[s(t)f_{sk} e^{j(\omega_C+\omega_I)t} + i(t)f_{ik} e^{j(\omega_C+\omega_I)t}] \quad (3a)$$

where $Re\{.\}$ denotes the real part of a signal, ω_C is the nominal frequency of the first mixer, and ω_I denote the frequency of the second local oscillator. The multiplication by 2 is introduced for convenience.

Equation (3a) can be rewritten as:

$$r_k(t) = s(t)f_{sk} e^{j(\omega_C+\omega_I)t} + s^*(t) f^*_{sk} e^{-j(\omega_C+\omega_I)t} + i(t)f_{ik} e^{j(\omega_C+\omega_I)t} + i^*(t)f^*_{ik} e^{-j(\omega_C+\omega_I)t} \quad (3b)$$

where * denotes complex conjugate.

The first mixer signal is expressed as:

$$x_{LO}(t)=2\cos(\omega_C t+\theta)= e^{-j(\omega_C t+\theta)} + e^{j(\omega_C t+\theta)} \quad (4)$$

where θ is the phase offset. Thus, no phase synchronization is assumed.

After the first mixer, the signals are downconverted to IF stage. Then Bandpass Filters (BPF's) with the center frequency of ω_I is employed to select the channel and suppress the high frequency components. The output of the BPF's is:

$$r_{IF,k}(t) = s(t)f_{sk} e^{j(\omega_I t-\theta)} + s^*(t)f^*_{sk} e^{-j(\omega_I t-\theta)} +$$

$$i(t)f_{ik}e^{j(\omega_I t-\theta)} + i^*(t)f^*_{ik}e^{-j(\omega_I t-\theta)} \quad (5)$$

Finally, the IF signals are further downconverted to baseband and processed by lowpass filters. At this point, A/D conversion is performed. The signal observation corresponding to the kth antenna $X_k(n)$ is given by:

$$X_k(n) = \mathrm{Re}\{s(n)f_{sk}e^{-j\theta} + i(n)f_{ik}e^{-j\theta}\} \quad (6)$$

where n represents the discrete time index.

For BPSK signals, since $s(t)$ and $i(t)$ are real-valued, so (6) can be written as:

$$X_k(n) = a_k s(n) + b_k i(n) \quad (7)$$

Where $a_k = \mathrm{Re}\{f_{sk}e^{-j\theta}\}$ and $b_k = \mathrm{Re}\{f_{ik}e^{-j\theta}\}$.

Usually the signals are each processed in frames of length N. Thus, while $s(n)$, $i(n)$ and $X_k(n)$ in (7) are one-sample signals, s_N, i_N and $X_{N,k}$ are used to denote blocks of signals, each containing N successive samples. Hence,

$$X_{N,k} = a_k s_N + b_k i_N \quad (8)$$

Therefore, the signal observation matrix is expressed as:

$$\mathbf{X} = \begin{bmatrix} X_{N,1} \\ X_{N,2} \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} s_N \\ i_N \end{bmatrix} = \mathbf{AS} \quad (9)$$

In this system model, **X** is the 2 by N observation matrix, **A** is the unknown 2 by 2 mixing matrix, and **S** is the 2 by N source signal matrix. ICA will be adopted to recover **S** given **X**, based on the assumption that the desired and interfering signals are statistically independent. This is achieved by identifying a 2 by 2 demixing matrix **W**, so that the components in the product **WX** will be as independent as possible.

The successful ICA separation requires the non-singularity of the mixing matrix **A**. This is guaranteed by the randomness of the channel's fading coefficients.

ICA has the inherent problem of order ambiguity, which means prior information is needed to properly identify the desired signal among the extracted signals. Thus, reference sequences are necessary. In most communication standards, this condition is satisfied [7].

3 Time-varying Channels

There are two types of learning algorithms: block adaptation and sequential adaptation.

Block algorithms use a block of data to establish statistical properties. Specifically, the "expectation" operator is estimated by the average over L data points, where L is the block size. The performance is improved when the estimation is more accurate, i.e., L is larger. However, it is very important that the mixing matrix stays approximately constant within one processing block.

On the other hand, the online sequential algorithms, in which update occurs once for every received sample, can better track the time variation. A difficult task required by the online algorithms is a proper choice of the learning rate, which is typically carried out by trial and error. Besides online gradient ICA, the well-known online ICA techniques also include EASI (Equivariant Adaptive Separation via Independence) and natural gradient ICA.

Fast-ICA [8] is a classic block ICA algorithm. It is highly efficient for relatively stationary channels, but it fails to converge for rapidly time-varying conditions. Recently we proposed a new Optimum Block Adaptive ICA (OBA/ICA) algorithm that exhibits significantly better convergence properties in dynamic conditions [9]. However, a new online ICA algorithm is still desirable for time-varying channel conditions, because the block algorithms cannot achieve good performance for small block sizes.

In the following, the online ICA algorithm with Individual Adaptation (IA-ICA) is presented as a special case of OBA/ICA. It uses individualized learning rates to perform online adaptation. The choice of the learning rate is no longer needed.

4 IA-ICA Algorithms for Signal Separation in Dynamic Conditions

In this section, the OBA/ICA is briefly introduced first, and online IA-ICA is then obtained as a special case of OBA/ICA.

j: iteration index.

M: number of observations.

L: length of the processing block.

$\underline{w(j)} = [w_1(j)\ w_2(j)\ \dots\dots\ w_M(j)]^T$: the current row of the separation matrix for the jth iteration. ($i = 1, 2, \dots, M$)

$x_{l,i}(j)$: the ith signal in the lth observation data vector for the jth iteration. ($l = 1, 2, \dots, L$)

$\underline{X_l(j)} = [x_{l,1}(j)\ x_{l,2}(j)\ \ldots\ldots\ x_{l,M}(j)]^T$: lth signal observation for the jth iteration.

$[G]_j = \left[\underline{X_1(j)} \quad \underline{X_2(j)} \quad \ldots.. \quad \underline{X_L(j)}\right]^T$: Observation matrix for the jth iteration.

Kurtosis is used as the measure of statistical independence. The lth kurtosis value for the jth iteration is expressed according to the definition as

$$kurt_l(j) = \mathrm{E}\{[\underline{w^T(j)}\ \underline{X_l(j)}]^4\} - 3 \qquad (10)$$

where it is assumed that the signals and $\underline{w(j)}$ both have been normalized to unit variance.

Then, the kurtosis vector for the jth iteration is given by

$$\underline{kurt(j)} = [kurt_1(j)\ kurt_2(j) \ldots\ldots\ kurt_L(j)]^T \qquad (11)$$

Now the updating formula with individual learning rates can be written as

$$\underline{w(j+1)} = \underline{w(j)} + [MU]_j \nabla_B(j) \qquad (12)$$

where

$$\nabla_B(j) = \frac{\partial\{\underline{kurt^T(j)kurt(j)}\}}{\partial \underline{w(j)}} = \frac{1}{L}\left[\frac{\partial\{\underline{kurt^T(j)kurt(j)}\}}{\partial w_1(j)} \ldots \frac{\partial\{\underline{kurt^T(j)kurt(j)}\}}{\partial w_M(j)}\right]^T \qquad (13)$$

and

$$[MU]_j = \begin{bmatrix} \mu_{B1}(j) & \ldots.. & 0 \\ \ldots.. & \ldots.. & \ldots.. \\ 0 & \ldots.. & \mu_{BM}(j) \end{bmatrix} \qquad (14)$$

Note that in (12), the weights are updated in the direction of the gradients, because the performance function is to be maximized.

On the other hand, the Taylor's series expansion for the lth kurtosis value in the $(j+1)^{th}$ iteration is

$$kurt_l(j+1) = kurt_l(j) + \sum_{i=1}^{M} \frac{\partial kurt_l(j)}{\partial w_i(j)} \Delta w_i(j) + \ldots \qquad (15)$$

Where

$$\Delta w_i(j) = w_i(j+1) - w_i(j)$$

In practice, if $\Delta w_i(j)$ is confined to be small enough, the higher order derivative terms in (15) can be omitted.

Equation (15) can be written for every l in a matrix-vector form. Combining the resulting equation and (12), the following is obtained after some derivation:

$$\underline{kurt(j+1)} = \underline{kurt(j)} + \frac{32}{L}[C]_j^3[G]_j[MU]_j[G]_j^T[C]_j^3\underline{kurt(j)} \qquad (16)$$

To identify the optimum $[MU]_j$, the total squared kurtosis $\underline{kurt^T(j)kurt(j)}$ is to be maximized. Thus, the following condition should be met:

$$\frac{\partial\{\underline{kurt^T(j+1)kurt(j+1)}\}}{\partial \mu_{Bi}(j)} = 0$$
$$i = 1,2,\ldots, M \qquad (17)$$

The optimum learning rates $\mu_{Bi}(j)$ are determined by substituting (16) into (17). After tedious derivation, the update equation (12) employing the optimum $[MU]_j$ is obtained as

$$\underline{w(j+1)} = \underline{w(j)} - 0.25[R]_j^{-1}\underline{q(j)} \qquad (18)$$

Where $\underline{q(j)}$ and $[R]_j$ are defined as

$$\underline{q(j)} = [G]_j^T[C]_j^3\underline{kurt(j)} = [q_1(j)\ldots q_M(j)]^T \qquad (19)$$

$$[R]_j = [G]_j^T[C]_j^6[G]_j = [R_{mn}(j)]$$
$$1 \le m,\ n \le M \qquad (20)$$

Equation (18) is the OBA/ICA algorithm.

To obtain IA-ICA, we assume block size L=1, so the block adaptation reduces to a sequential one. In this case, $[R]_j$ and $\underline{q(j)}$ defined in (19) and (20) becomes:

$$[R]_j = \underline{X(j)}[\underline{w^T(j)X(j)}]^6\underline{X^T(j)} \qquad (21)$$

$$\underline{q(j)} = \underline{X(j)}[\underline{w^T(j)X(j)}]^3 kurt(j) \qquad (22)$$

Note that in (22), $l = 1, 2, \ldots, L$ (8) the kurtosis vector has degenerated to a scalar.

From (21)

$i = 1, 2, \ldots, M$ (9)

$$[R]_j^{-1} = \frac{[\underline{X(j)X^T(j)}]^{-1}}{[\underline{w^T(j)X(j)}]^6} \qquad (23)$$

Substitute (22) and (23) into (18), the IA-ICA is obtained:

$$\underline{w}(j+1) = \underline{w}(j) - 0.25kurt(j)\frac{[\underline{X}(j)\underline{X}^T(j)]^{-1}\underline{X}(j)}{[\underline{w}^T(j)\underline{X}(j)]^3} \quad (24)$$

5 Simulation Results

To study the performance of the proposed technique, computer simulations are performed. Two sequential ICA algorithms, IA-ICA and online gradient ICA [10, pp177], are used to find the demixing matrix. During the adaptation process, the data block is repeated if necessary. The converged demixing matrix is applied to the data block to recover the source signals.

Since the wireless channel is assumed to be time-varying within processing blocks, small block size L has to be adopted. Also, the choice of the block size should be made according to the data rate and the receiver's processing speed. In our experiment, L=100.

The time-varying mixing matrix is randomly chosen as:

$$A = \begin{bmatrix} 10 + l\Delta & 5 \\ 7 & 20 + l\Delta \end{bmatrix} \quad (25)$$

where $l = 1, 2, \dots, L$, and Δ is the parameter reflecting the speed of channel variation. Here, it is assumed that the channel's transfer function is frequency-flat over the signal band. Also, the sampling interval of the receiver's A/D converter is negligible compared with $1/\Delta$, which represents the rate of the channel's time variation.

The separation performance is measured by Signal-to-Interference Ratio (SIR) defined as:

$$\text{SIR} = 10\log_{10}\left(\frac{1}{L}\sum_{k=1}^{L}\frac{s(k)^2}{[s(k)-y(k)]^2}\right) \quad (26)$$

where $s(k)$ is the kth sample of the desired signal, $y(k)$ is the estimate of the $s(k)$ obtained at the output of the demodulation stage. SIR represents the average ratio of the desired signal power to the power of the estimation error.

In our simulations, the time-variation parameter Δ is varied from 0.01 to 0.1. This represents about 10% to 100% of variation in the mixing matrix's coefficients within the same block. For each Δ value, 100 Monte Carlo simulations are performed to get the average performance. When online gradient ICA is used, the proper learning rate has to be chosen by trial and error. It is found that the suitable learning rate is about 0.1 under our experiment setup.

The average SIR and the number of iterations required for convergence adopting IA-ICA and online gradient ICA are shown in Figs. 2 and 3. It is seen from Fig. 2 that IA-ICA achieves better separation performance than online gradient ICA. In Fig. 3, it is shown that IA-ICA converges faster if the time-variation parameter Δ is less than 0.09.

To study the performance of the new technique under noisy conditions, thermal noise is added to the received signal. For this set of simulations, $\Delta = 0.01$. Instead of (25), the separation performance is expressed as the Bit-Error-Rate (BER) versus the input Signal to Noise Ratio (SNR) in the range of 0 to 10dB. Figure 4 plots the results obtained from IA-ICA and online gradient ICA. It is seen that both ICA algorithms achieve similar BER, and IA-ICA's performance is slightly better.

6 Conclusions

In this paper, a novel sequential ICA algorithm with individual adaptation (IA-ICA) is proposed for signal separation in dynamic channel conditions. The algorithm automatically selects an individualized set of learning rates, thus achieves better performance than the traditional online gradient ICA. The new technique is applied to interference suppression in BPSK wireless receivers. Computer simulation results illustrate the advantages of IA-ICA. Also, it is shown that sequential ICA algorithms are capable of performing signal separation under noisy conditions. In the future, extension of IA-ICA to complex-valued signals will be investigated.

References:

[1] E. Buracchini, The Software Radio Concept, *IEEE Communications Magazine*, September 2000.

[2] Jack P.F. Glas, Digital I/Q imbalance compensation in a low-IF receiver, *Global Telecommunications Conference*, November 1998.

[3] I. Kostanic and W. Mikhael, Blind source separation technique for the reduction of co-channel interference, *IEE Electronics Letters*, Vol. 38, No. 20, September 2002.

[4] T. Yang, W. Mikhael, A General Approach for Image and Co-channel Interference Suppression in Diversity Wireless Receivers Employing ICA, *Journal of Circuits, Systems, and Signal Processing (CSSP)*, VOL. 24, No. 3, pp. 317-327, September 2004.

[5] T. Yang, W. Mikhael, A General Interference Suppression Scheme for Diversity Wireless Receivers, *Invited paper*, *Workshop on Wireless Circuits and Systems (WoWCAS)*, University of British Columbia, Vancouver, Canada, May 2004.

[6] Mikko Valkama, Markku Renfors, Visa Koivunen, Advanced methods for I/Q imbalance compensation in communication receivers, *IEEE Transactions On Signal Processing*, vol. 49, no. 10, Oct. 2001.

[7] T. Rappaport, Wireless Communications: Principles and Practice, Prentice Hall Inc., 1996.

[8] A. Hyvärinen and E. Oja, A fast fixed-point algorithm for independent component analysis, *Neural Computation*, 9(7): 1483-1492, 1997.

[9] W. Mikhael, T. Yang, Optimum Block Adaptive Algorithm for Gradient Based Independent Component Analysis (OBA/ICA) for Time Varying Wireless Channels, *Proceedings of The 62nd IEEE Vehicular Technology Conference*, Dallas, Texas, Sept. 2005.

[10] A. Hyvarienen, J. Karhunen, E. Oja, Independent Component Analysis, John Wiley and Sons, 2001.

Fig. 1 Diversity BPSK receiver structure

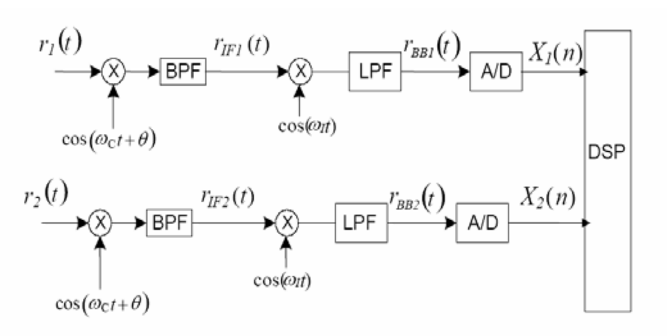

Fig. 2 Average SIR achieved by IA-ICA and online gradient ICA

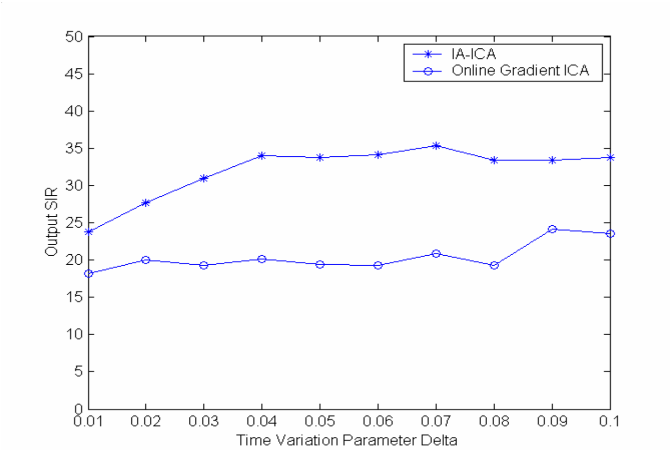

Fig. 3 Average number of iteration to convergence adopting IA-ICA and online gradient ICA

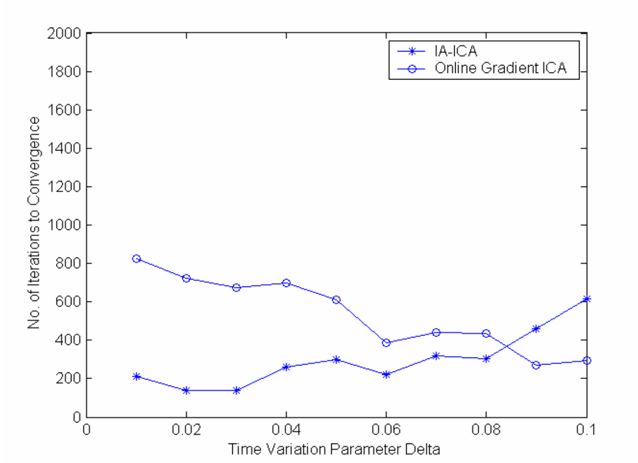

Fig. 4 BER performance achieved by IA-ICA and online gradient ICA under noisy conditions