**Dariusz PUCHALA, Mykhaylo YATSYMIRSKYY**
Institute of Information Technology, Technical University of Lodz

# Fast parametrized biorthogonal transforms

***Abstract.*** *In this paper the authors propose the buildling scheme of fast parametrized biorthogonal transforms taking advantage of data-flow graphs for calculation of fast orthonormal transforms and two-point biorthogonal butterfly operators.*

***Streszczenie.*** *W pracy zaproponowano schemat budowy szybkich parametryzowanych przekształceń biortogonalnych oparty o diagramy przepływowe dla szybkich przekształceń ortonormalnych i dwupunktowe biortogonalne operatory motylkowe. (***Szybkie parametryzowane transformacje bi-ortogonalne***)*

**Keywords:** fast parametrized linear transforms
**Słowa kluczowe:** szybkie parametryzowane przekształcenia liniowe

## Introduction

In latter years, we could observe a dynamic development of techniques for synthesis of parametrized linear orthonormal transforms with properties that enable construction of effective computational algorithms. Parametrized transforms in contrast to transforms with fixed base vectors have the fundamental advantage of the ability to adapt to statistical characteristics of signals. For example, fast parametric Haar-type transforms that allow to set the first base vector to any *a priori* assumed form were proposed in [1]. In paper [2] fast two-stage parametric transform allowing to modify all base vectors by manipulating the number of $N/2\log_2 N$ individual parameters was proposed and the steepest decent gradient technique was applied in parameters adaptation process. In papers [3] and [4] parametrized variants of well known discrete Fourier, Hartley and Slant-Hadamard transforms were formulated.

In this contribution authors propose the building scheme of fast parametrized biorthogonal transforms (FPBT) that takes advantage of biorthogonal butterfly operators and the data-flow graphs of fast algorithms for computation of orthonormal transforms.

## Fast parametrized biorthogonal transforms

In the proposed building scheme of FPBT we take as the starting point a data-flow graph of fast algorithm of orthonormal transform with an arbitrary structure. In the remaning part of the paper we make use of the following exemplary two-stage graph presented in Figure 1. Its structure is characterized by a symmetry facilitating software and hardware implementation of the algorithm and relatively large number of butterfly operators that increases the capacity and adaptive capabilities of the structure. In case of an orthonormal transform butterfly operators "•" represent selected affine operations that can be described in matrix form i.a. as $A_{ij}$,

$$A_{ij} \triangleq \begin{bmatrix} c_{ij} & s_{ij} \\ -s_{ij} & c_{ij} \end{bmatrix}, \quad A_{ij}^{-1} \triangleq \begin{bmatrix} c_{ij} & -s_{ij} \\ s_{ij} & c_{ij} \end{bmatrix},$$

where $c_{ij} \equiv \cos(\alpha_{ij})$ and $s_{ij} \equiv \sin(\alpha_{ij})$, $A_{ij}^{-1} = A_{ij}^T$ is an inverse operator and $\alpha_{ij}$ is a parameter of an operator (see [2]). As a result we obtain two-stage transformation with the number of $N(\log_2 N - 0.5)$ parameters.

In order to obtain biorthogonal transformation we replace operators $A_{ij}$ with biorthogonal operators $B_{ij}$ proposed in paper [5]. Then inverse butterfly operators are formulated as $B_{ij}^{-1}$ (see paper [5]).

$$B_{ij} \triangleq \begin{bmatrix} a_{ij} & b_{ij} \\ c_{ij} & d_{ij} \end{bmatrix}, B_{ij}^{-1} \triangleq \frac{1}{\det(B_{ij})} \begin{bmatrix} d_{ij} & -b_{ij} \\ -c_{ij} & a_{ij} \end{bmatrix}.$$

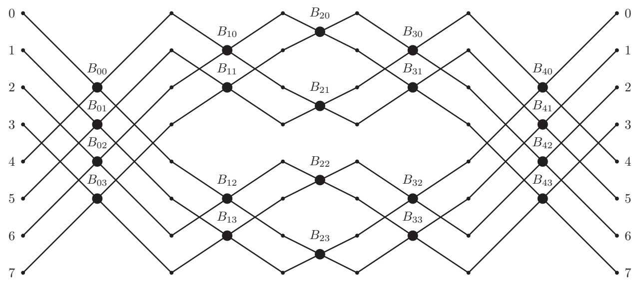


Fig. 1. Two-stage structure of fast parametrized biorthogonal transform for $N = 8$

The following condition $\det(B_{ij}) \neq 0$ must be satisfied to ensure the invertibility of operators. Then forward transform $U$ is constructed as a product of $U_i$ matrices that hold butterfly operators $B_{ij}$ belonging to succesive stages, i.e. $U = U_{n-1} \ldots U_1 U_0$, where $i$ is a stage index and $j$ is an index of operator at $i$-th stage. Inverse transformation $U^{-1}$ is obtained as a product of inverse matrices $U_i^{-1}$ taken in reverse order, i.e. $U^{-1} = U_0^{-1} U_1^{-1} \ldots U_{n-1}^{-1}$ where $U_i^{-1}$ matrices hold inverse biorthogonal operators $B_{ij}^{-1}$. As an effect of application of biorthogonal operators the norms and the angles between base vectors of $U$ transform can take any values always different to zero. Both transforms $U$ and $U^{-1}$ hold dual bases in $\mathcal{R}^N$ space and ensure perfect reconstruction of signals. Hence, we name $U$ transform the *fast biorthogonal transformation*. It should be noted that each operator $B_{ij}$ is described with four adjustable parameters $a_{ij}, b_{ij}, c_{ij}$ and $d_{ij}$ what finally results in the adaptable *parametrized* transform. For the considered two-stage structure an entire number of parameters equals $4N(\log_2 N - 0.5)$.

Biorthogonal operators $B_{ij}$ are not convinient for automatic adaptation, e.g. with gradient decent techniques, since $\det(B_{ij}) \neq 0$ condition must be satisfied. Moreover, what is to be proved in following part of this section, the number of $4N(\log_2 N - 0.5)$ parameters is redundant and can be reduced to $2N\log_2 N$ with $U$ transform undisturbed if only the additional condition $a_{ij} \neq 0$ is satisfied for each pair of indices $i$ and $j$. The excessive number of parameters is undesired since it increases the dimensionality of a search space during the adaptation process and induces longer description of transformation itself.

It is to be noted that with assumed conditions each $B_{ij}$ operator can be expressed in the following equivalent form:

$$B_{ij} = \begin{bmatrix} p_{ij} & 0 \\ 0 & q_{ij} \end{bmatrix} \cdot \begin{bmatrix} 1 & u_{ij} \\ v_{ij} & 1 + u_{ij}v_{ij} \end{bmatrix},$$

where $p_{ij} = a_{ij}$, $q_{ij} = \det(B_{ij})/a_{ij}$ and $u_{ij} = b_{ij}/a_{ij}$,

$v_{ij} = a_{ij}c_{ij}/\det(B_{ij})$, while $p_{ij},\ q_{ij} \neq 0$. Multipliers $p_{ij}$ and $q_{ij}$ are only the scaling factors for outputs of an operator and that mutliplications can be enclosed within operators $B_{(i+1)j}$ in the consecutive layer. Let $r_{ij}$ and $t_{ij}$ be the multipliers transferred from $(i-1)$-th layer to operator $B_{ij}$. Evidently $r_{ij}, t_{ij} \neq 0$ for each pair of $i, j$. Then:

$$B_{ij}\begin{bmatrix} r_{ij} & 0 \\ 0 & t_{ij} \end{bmatrix} = \begin{bmatrix} \hat{p}_{ij} & 0 \\ 0 & \hat{q}_{ij} \end{bmatrix} \cdot \begin{bmatrix} 1 & \hat{u}_{ij} \\ \hat{v}_{ij} & 1+\hat{u}_{ij}\hat{v}_{ij} \end{bmatrix},$$

where $\hat{p}_{ij} = p_{ij}r_{ij}$ and $\hat{q}_{ij} = q_{ij}t_{ij}$, $\hat{u}_{ij} = u_{ij}(t_{ij}/r_{ij})$ and $\hat{v}_{ij} = v_{ij}(r_{ij}/t_{ij})$. Applying the above-cited scheme layer by layer we obtain a structure where $B_{ij}$ operators are replaced with operators of the following form:

$$C_{ij} \triangleq \begin{bmatrix} 1 & \hat{u}_{ij} \\ \hat{v}_{ij} & 1+\hat{u}_{ij}\hat{v}_{ij} \end{bmatrix}.$$

Since for each pair of indices $i, j$ the condition $\det(C_{ij}) = 1$ is satisfied an operator inverse to $C_{ij}$ takes form:

$$C_{ij}^{-1} \triangleq \begin{bmatrix} 1+\hat{u}_{ij}\hat{v}_{ij} & -\hat{u}_{ij} \\ -\hat{v}_{ij} & 1 \end{bmatrix}.$$

Multipliers $\hat{p}_{ij}$, $\hat{q}_{ij}$ emerged from the last layer form an additional diagonal layer with elements $d_i$ for $i = 0, 1, \ldots, N-1$. Hence, assuming that $\det(B_{ij}) \neq 0$ and $a_{ij} \neq 0$ for all $i, j$ we obtain structure (see Fig. 2) equivalent to one presented in Figure 1. However, such structure has the following advantages: firstly each operator $C_{ij}$ ($C_{ij}^{-1}$) regardless of the values of parameters has its determinant equal to one, secondly operators $C_{ij}$ ($C_{ij}^{-1}$) require only two parameters and for calculation of their outputs only two multiplications are needed (see (1)). That is twice less number of multiplications than in case of $B_{ij}$ operators while a constant number of additions is retained.

$$(1)\qquad C_{ij} = \begin{bmatrix} 1 & 0 \\ \hat{v}_{ij} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \hat{u}_{ij} \\ 0 & 1 \end{bmatrix},$$

$$C_{ij}^{-1} = \begin{bmatrix} \hat{u}_{ij} & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \hat{v}_{ij} & -1 \\ 1 & 0 \end{bmatrix}.$$

The structure of FPBT from Figure 2 can be characterized by the computation complexity: $L_{MUL} = 2N\log_2 N$ and $L_{ADD} = 2N(\log_2 N - 0.5)$ where $L_{MUL}$ and $L_{ADD}$ are the numbers of floating-point multiplications and additions respectively.

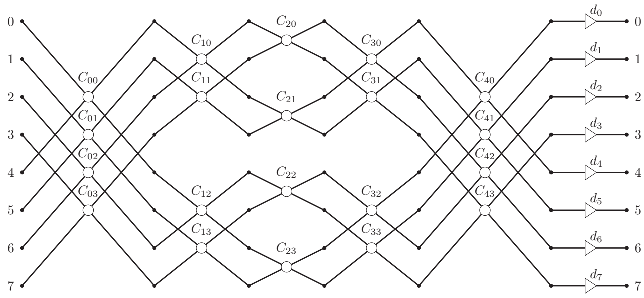

Fig. 2. Equivalent fast structure with the reduced number of parameters for biorthogonal transform of $N = 8$ length

**Experimental results**

In order to verify the effectiveness of FPBT with structure from Figure 2 several experiments in Generalized Wiener filtering (GWF) [6] were conducted. The scheme of GWF is presented in Figure 3. At its input we feed $x$ signal buried in additive noise $\epsilon$ which is filtered in $A$ transform domain with filter $G$. At the output a reconstructed with transform $B$ version $\hat{x}$ of input signal is obtained. The practical selection criteria of matrices $A$, $B$ and $G$ should take into consideretion not only the minimization of mean square error (MSE) between demanded $x$ and the reconstructed signal $\hat{x}$ but also the minimization of the number of computations. Hence, in practical solutions $A$ represents usually a fast orthonormal transform, $B = A^T$ and $G$ is taken as a diagonal matrix $D_G$.

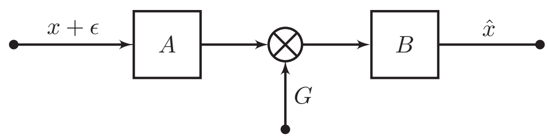


Fig. 3. Generalized Wiener filtering

In the first part of experiments we considered electrocardiogram signals (ECG) buried in pink noise i.e. noise with power spectra density inversely proportional to frequency. The results in filtering with FPBT were confronted with results for discrete cosine transform (DCT) and optimally selected diagonal filter matrix $D_G$. The choice of DCT was motivated by a fact that it produced best results among known fast orthonormal transforms with fixed bases. In FPBT filtering scheme we resigned from matrix $G$ in explicit form connecting two structures from Figure 2 where one structure is transposed. The resulting data-flow diagram represents product of two matrices $A$ and $B$ and from a structural viewpoint it can be described as two-stage structure - diagonal layer - two-stage structure. Input vectors $x+\epsilon$ of the following lenghts $N = 32, 64$ and $128$ (see Figure 4) were considered during this experiment. Moreover, to gain the statistical significance we examined $60$ sets with $5000$ vectors each. However, calculation of optimal filtering matrix for DCT as well as selection of parameters for FPBT with use of steepest decent technique were realized on the basis of a single set of training vectors. The representative results for FPBT and DCT based filtering are presented in Table 1 in the form of peak signal to noise ratios (PSNR).

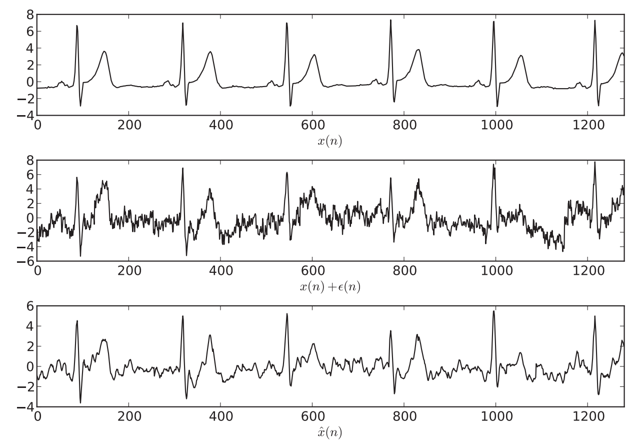

Fig. 4. Exemplary portions of ECG signals: demanded $x$, input $x+\epsilon$ and filtered $\hat{x}$

An analysis of obtained results exhibits significant advantage of FPBT over scheme based on DCT and diagonal filtering matrix. The mean value of PSNR gain is highest for input vectors with length $N = 64$ and it equals $1.33$ dB. For the remaning values of $N = 32$ and $128$ the resulting gain is $1.09$ dB and $0.74$ dB high respectively. The lowest advan-

Table 1. The PSNR results in filtering of all sets of $5000$ vectors with lengths $N = 32$, $64$ and $128$ with use of DCT, FPBT scheme and optimal filter (OF) where min., mean and max. symbolize minimal, mean and maximal values respectively

| | DCT [dB] | | | FPBT [dB] | | | OF [dB] |
|---|---|---|---|---|---|---|---|
| $N =$ | min. | mean | max. | min. | mean | max. | mean |
| $32$ | $20.9$ | $23.2$ | $25.5$ | $21.6$ | $24.3$ | $26.9$ | $24.5$ |
| $64$ | $22.1$ | $23.9$ | $26.0$ | $23.4$ | $25.3$ | $27.5$ | $25.6$ |
| $128$ | $25.2$ | $26.9$ | $28.2$ | $26.7$ | $27.6$ | $28.1$ | $29.0$ |

tage of FPBT for $N = 128$ might be a result of convergence of gradient technique to local minima. It should be noted that the mean distance between FPBT and optimal filtering for the considered values of $N$ equals in sequence $0.23$ dB, $0.29$ dB and $1.43$ dB. For $N = 32$ and $64$ the distance is small and it can be stated that FPBT results are close to optimal.

In DCT based filtering scheme optimal $2m + 1$ diagonal filtering matrices $D_G$ for $m = 1, 2, \ldots, N-1$ were also considered. The obtained results revealed that the values of $m$ close to $N$ were required to obtain PSNR values comparable with those of FPBT. However, high values of $m$ induce significant increase of computational complexity. For instance for $N = 64$ the sole filtering required additional $3958$ multiplications and $3894$ additions while FPBT scheme demanded the total number of $1472$ multiplications and $1408$ additions.

In the second part of experiments we confronted the proposed FPBT based filtering scheme with GWF scheme taking advantage of fast parametrized Slant-Hadamard transform (FPSHT) formulated in paper [4]. The experiment complied with the methodology proposed in [4], i.e. input vectors can be characterized by general autocorrelation model (2) and are buried in white uncorrelated noise keeping the unit value of signal to noise ratio. The obtained results in MSE are collected in Table 2. The autocorrelation matrix for demanded vectors $x$ takes the following form:

$$R_x(m,n) = \exp\left\{-\left[(\alpha m^{r_1})^h + (\chi n^{r_2})^h\right]^{1/h}\right\} \quad (2)$$

for $n$, $m = 0, 1, \ldots, N-1$ where: $r_1 = 1.137$, $r_2 = 1.09$, $h = \sqrt{2}$, $\alpha = 0.025$ i $\chi = 0.019$.

Table 2. The MSE results obtained for general autocorrelation model $R_x$ and DCT, FPBT, fast parametrized Slant-Hadamard transform (FPSHT) and optimal filtering (OF) scheme

| | $N$ | | | | |
|---|---|---|---|---|---|
| | $4$ | $8.$ | $16$ | $32$ | $64$ |
| OF | $0.2134$ | $0.1342$ | $0.0942$ | $0.0809$ | $0.0814$ |
| DCT | $0.2139$ | $0.1364$ | $0.1024$ | $0.0995$ | $0.1181$ |
| FPSHT | $0.2139$ | $0.1364$ | $0.1023$ | $0.0985$ | $0.1149$ |
| FPBT | $0.2134$ | $0.1342$ | $0.0942$ | $0.0809$ | $0.0814$ |

It should be noted that FPSHT based scheme produced results significantly better than DCT filtering. However, the results obtained with FPBT are identical to the optimal values for all lengths of $N = 4$, $8$, $16$, $32$ and $64$. Hence, it can be stated that the proposed FPBT under the conditions dictated by the experiment is optimal in the sense of MSE metric keeping the fast computational structure at the same time.

## Conclusions

Biorthogonal linear transforms include the subclass of orthonormal transformations without restriction in the form of orthogonality and unit norms of base vectors. Hence, the possible range of their applications is much wider and includes directly such issues as: signal classification, generalized Wiener filtering [6] and blind source separation exploiting independent component analysis techniques [7].

The results obtained in this paper for general Wiener filtering of ECG signals corrupted with pink noise indicate a significant advantage of fast parametrized biorthogonal transforms which produced results better on average at about $1$ dB than filtering sheme using fast orthonormal DCT transform and diagonal or $2m + 1$ diagonal filtering matrix for $m = 1, 2, \ldots, N-1$. Also, the results obtained in the second part of experiment for signals with $R_x$ autocorrelation matrix (see (2)) indicate an advantage of FPBT not only over DCT but also over fast parametrized Slant-Hadamard transform. Equally important is the fact that they do not differ from optimal results. Thus proving the effectiveness of the proposed biorthogonal transformation in signal filtering we indicate as future directions of research a verification of the possibilities to use the proposed transforms in other, above-mentioned in this contribution, issues in the field of analysis and filtering of signals.

It should be noted that proposed in this paper method of reduction of the number of multiplication in biorthogonal structure can be successfully applied to reduce the complexity of other fast structures, including the structures for calculation of orthonormal transforms. A similar method of tangent multipliers with application to two-stage fast computational algorithms for cosine and sine transforms of type two and three was proposed in paper [8].

## REFERENCES


[1] S. Minasyan, J. Astola, D. Guevorkian, *On a Class of Parametric Transforms and Its Application to Image Compression*, EURASIP Journal on Advances in Signal Processing, 2007.

[2] D. Puchala, M. Yatsymirskyy, *Fast neural networks learning techniques for signal compression*, Electrical Review, vol. 1, pp. 189-191, January 2010.

[3] S.Bouguezel, M.O.Ahmad, M.N.S.Swamy, *New Parametric Discrete Fourier and Hartley Transforms, and Algorithms for Fast Computation*, IEEE Trans. On Circuits And Systems, vol. 58, no. 3, pp. 562-575, March 2011.

[4] S.Agaian, K.Tourshan, J.P.Noonan, *Parametric Slant–Hadamard Transforms With Applications*, IEEE Signal Processing Letters, vol. 9, no. 11, November 2002.

[5] M. Yatsymirskyy, *Effective lattice wavelet transform*, Information Extraction and Process. 2011. Issue 34 (110), p.95-100.

[6] W.K.Pratt, *Generalized Wiener Filtering Computation Techniques*, IEEE Trans. On Computers, vol. C-21, pp. 636-641, July 1972.

[7] P.Comon, Ch.Jutten, *Handbook of Blind Source Separation. Independent Component Analysis and Applications*, Elsevier, 2010.

[8] M. Yatsymirskyy, *Uniform algorithms for fast computation of cosine and sine transform of the second and third form with tangent factors*, Automatyka AGH, vol. 3, pp. 727-741, 2003 (in polish).



***Authors:*** Ph.D. Dariusz Puchala, Prof. Mykhaylo Yatsymirskyy, Institute of Information Technology, Faculty of Technical Physics, Computer Science and Applied Mathematics, Technical University of Lodz, Wolczanska Str. 215, 90-924 Lodz, email: (dariusz.puchala, mykhaylo.yatsymirskyy)@p.lodz.pl