Dariusz PUCHALA, Mykhaylo YATSYMIRSKYY

Institute of Information Technology, Technical University of Lodz

Fast Parametrized Biorthogonal Transforms With Normalized Basis Vectors

Abstract. *In this paper the authors propose a scheme for constructing fast parametrized biorthogonal transforms with normalized, i.e. of unit length, basis vectors. As a result the transformation with basis vectors distributed over the surface of hypersphere with unit radius is obtained where the basis vectors need not to be mutually orthogonal.*

Streszczenie. *W pracy zaproponowano schemat budowy szybkich parametryzowanych przekształceń biortogonalnych o znormalizowanych, tj. posiadających długości jednostkowe, wektorach bazowych. W rezultacie otrzymujemy przekształcenie o wektorach bazowych rozłożonych na powierzchni hipersfery o promieniu jednostkowym, przy czym wektory bazowe nie muszą być wzajemnie prostopadłe. (***Szybkie parametryczne przekształcenia biortogonalne o znormalizowanych wektorach bazowych***)*

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

Introduction

The parametrized transforms are among most dynamically developing tools for processing and analysis of digital signals. For this reason, in the last decade, a number of papers devoted to the synthesis of fast parametrized transforms have emerged, including, e.g. Haar type transforms [1], two-stage transformations [2], different variants of Fourier, Hartley [3] and Slant-Hadamard [4] transforms. The advantage of parametrized transforms is that they can be adapted to the statistical characteristics of signals. It should be noted that all of the aforementioned transformations are orthonormal.

In paper [5] a scheme for constructing fast biorthogonal parametrized transforms was proposed. Their advantage over orthonormal transforms is the lack of restrictions of unit length and perpendicularity of base vectors. This translates into greater flexibility and wider possibilities of their direct use in practice, including the generalized Wiener filtering [6], classification of signals and separation of signals based on independent component analysis [7]. The absence of those two restrictions may, however, in certain applications, e.g. in signal classification, increase the dimensionality of the search space, which may lead to longer times of adaptation, and even to worse final results.

In this paper we propose a two-stage scheme of radix-2 type for synthesis of fast parametrized biorthogonal transforms with normalization (FPBNT), i.e. with only one restriction on unit length of basis vectors. The base vectors are then distributed over the surface of the hypersphere of unit radius, and the lack of orthogonality restriction makes the angles between base vectors not necessarily right.

Biorthogonal transforms with normalized base vectors

The synthesis of FPBNT with two-stage structure of radix-2 type is based on the following observations, which are presented below without proofs:

1. if we are given a set $\{u_i\}$ of M vectors determining the subspace of $\mathcal{R}^N$ and their linear combination of the form: $y = \sum_{i=0}^{M-1} \gamma_i u_i$, then $\|y\|_2 = 1$ provided that $\sum_{i=0}^{M-1} \gamma_i^2 = 1$ and vectors $\{u_i\}$ are orthonormal,
2. if we are given two sets of vectors $\{u_i\}$ and $\{v_i\}$ determining mutually perpendicular subspaces in $\mathcal{R}^N$, then the subspaces determined by any linear combinations of vectors $\{u_i\}$ and $\{v_i\}$ respectively must be mutually perpendicular.

In the radix-2 scheme at each step two base vectors (resulting vectors) of transformation are created as linear combinations of two vectors from the previous step (input vectors). In the case of orthonormal transforms such combinations are calculated with rotation operators, e.g. of the form (1a). The rotations guarantee orthonormality of resulting vectors providing that input vectors are orthonormal. For biorthogonal transforms [5] operators take the general form (1b), which is described by four independent parameters. On the basis of the first observation we find that in order to obtain the unit length of the resulting vectors, while providing the orthogonality of input ones, it is enough to take such coefficients of the linear combination whose squares add up to one. This leads to the formulation of the operator of an exemplary form (1c), where the squared elements in each row of the matrix B_{ij} add up to one. It should be noted that the angle calculated between the resulting vectors need not to be right.

(1a) $$O_{ij} = \begin{bmatrix} \cos(\alpha_{ij}) & \sin(\alpha_{ij}) \\ -\sin(\alpha_{ij}) & \cos(\alpha_{ij}) \end{bmatrix},$$

(1b) $$A_{ij} = \begin{bmatrix} a_{ij} & b_{ij} \\ c_{ij} & d_{ij} \end{bmatrix},$$

(1c) $$B_{ij} = \begin{bmatrix} \cos(\alpha_{ij}) & \sin(\alpha_{ij}) \\ -\sin(\beta_{ij}) & \cos(\beta_{ij}) \end{bmatrix}.$$

Figure 1 shows a single step of calculation of FPBNT. The matrix of the first stage U_0 contains row vectors u_i for $i = 0, 1, \ldots, N-1$ of unit length, which are also pairwise orthogonal at the points of application of B_{0i} operators, i.e. $u_i \perp u_{i+N/2}$ for $i = 0, 1, \ldots N/2-1$. Matrix U_1 is a matrix of the second stage consisting of operators B_{0i}.

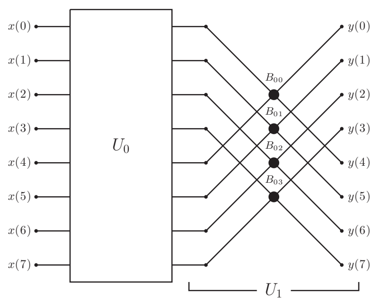

Fig. 1. Single step of calculation of $N = 8$ point FPBNT

The second observation allows to specify the conditions that must by met by row vectors of U_0 matrix to enable the insertion of the following steps for calculation of FPBNT. We consider the scheme of FPBNT with the second stage containing two computational steps (for $N = 8$ see Fig. 2). Based on the previous discussion we know that vectors $\hat{u}_i$ obtained after the first step, whose combinations are calcu-

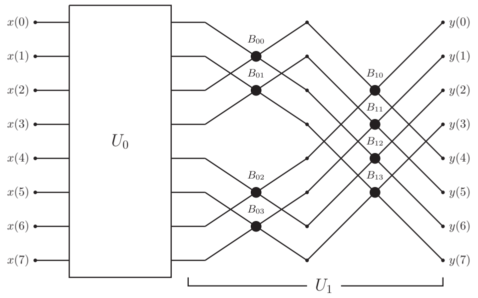

Fig. 2. Two steps of calculation of $N = 8$ point FPBNT

lated after that in the second step, should be pairwise perpendicular, i.e. $\hat{u}_i \perp \hat{u}_{i+N/2}$ for $i = 0, 1, \ldots N/2 - 1$, and should be of unit length. The vectors $\hat{u}_i$ are calculated as linear combinations of row vectors u_i of the first stage matrix U_0. Then, on the basis of the first observation we find that vectors u_i should be of unit length and vectors u_i, $u_{i+N/4}$, and also $u_{i+N/2}$, $u_{i+3N/4}$ for $i = 0, 1, \ldots, N/4 - 1$ should be pairwise perpendicular. This ensures the unit lengths of vectors $\hat{u}_i$ obtained after the first step. Then, based on the second observation, we conclude that in order to ensure the condition $\hat{u}_i \perp \hat{u}_{i+N/2}$ for $i = 0, 1, \ldots, N/2 - 1$ it is additionally required that the subspaces spanned respectively by the pairs of vectors u_i, $u_{i+N/4}$ and $u_{i+N/2}$, $u_{i+3N/4}$ be mutually perpendicular for $i = 0, 1, \ldots N/4 - 1$.

By applying the above reasoning recursively we get the conditions that must be met by the row vectors of U_0 matrix for the full number of $\log_2 N$ steps. Then it can be shown that U_0 matrix must be orthonormal. In Fig. (3) the full number of steps required for calculation of $N = 8$ point FPBNT is presented. The resulting transformation $V = U_1 U_0$, where U_0 is an orthonormal matrix (it can be a unit matrix, the matrix of fast parametrized orthonormal transform, etc.) and U_1 is the matrix of the second stage, is the transformation with row vectors of unit length for any values of operators coefficients.

It should be noted that in the considered example the permutations between steps of the second stage, i.e. points of application of operators B_{ij}, are only exemplary. Other permutations consistent with the radix-2 type scheme are also permissible.

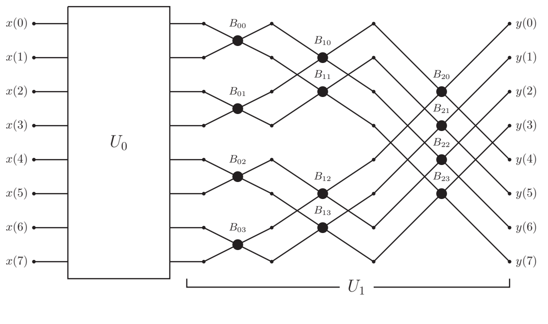

Fig. 3. The full number of steps of calculation of $N = 8$ point FPBNT

Equivalence with biorthogonal structure

In this section, we show that FPBNT with an additional diagonal matrix is equivalent to biorthogonal transform with the first stage orthonormal and second stage with identical structure but consisting of operators (1b). We use here the technique that is similar to one presented in paper [5] which allows to transfer between adjacent steps diagonal matrices that scale the outputs of individual operators. Let us consider the operator of the form (1b). This operator can be rewritten as:

$$A_{ij} = \begin{bmatrix} a_{ij} & b_{ij} \\ c_{ij} & d_{ij} \end{bmatrix} = \begin{bmatrix} p_{ij} & 0 \\ 0 & q_{ij} \end{bmatrix} \begin{bmatrix} \hat{a}_{ij} & \hat{b}_{ij} \\ \hat{c}_{ij} & \hat{d}_{ij} \end{bmatrix},$$

where $p_{ij} = \sqrt{a_{ij}^2 + b_{ij}^2}$, $q_{ij} = \sqrt{c_{ij}^2 + d_{ij}^2}$, $\hat{a}_{ij}$, $\hat{b}_{ij}$, $\hat{c}_{ij}$ and $\hat{d}_{ij}$ are the appropriate coefficients satisfying: $\hat{a}_{ij}^2 + \hat{b}_{ij}^2 = 1$, $\hat{c}_{ij}^2 + \hat{d}_{ij}^2 = 1$. If to the inputs of operator A_{ij} are connected operators from the previous step, which also exclude scaling matrices, then the scaling of the previous step can be implemented in the current step by taking into account the scaling factors in the structure of the operator A_{ij}, i.e.:

$$A_{ij} \begin{bmatrix} r_{ij} & 0 \\ 0 & t_{ij} \end{bmatrix} = \begin{bmatrix} a_{ij} r_{ij} & b_{ij} t_{ij} \\ c_{ij} r_{ij} & d_{ij} t_{ij} \end{bmatrix} = \\ = \begin{bmatrix} \hat{p}_{ij} & 0 \\ 0 & \hat{q}_{ij} \end{bmatrix} \begin{bmatrix} \bar{a}_{ij} & \bar{b}_{ij} \\ \bar{c}_{ij} & \bar{d}_{ij} \end{bmatrix},$$

where we have here: $\hat{p}_{ij} = \sqrt{(a_{ij} r_{ij})^2 + (b_{ij} t_{ij})^2}$ and $\hat{q}_{ij} = \sqrt{(c_{ij} r_{ij})^2 + (d_{ij} t_{ij})^2}$, $\bar{a}_{ij}$, $\bar{b}_{ij}$, $\bar{c}_{ij}$ and $\bar{d}_{ij}$ are the appropriate coefficients satisfying: $\bar{a}_{ij}^2 + \bar{b}_{ij}^2 = 1$ and $\bar{c}_{ij}^2 + \bar{d}_{ij}^2 = 1$. The scaling coefficients $\hat{p}_{ij}$ and $\hat{q}_{ij}$ excluded from the last step are transfered to the additional diagonal matrix D. In the whole structure we have then the operators of the form:

$$\bar{A}_{ij} = \begin{bmatrix} \bar{a}_{ij} & \bar{b}_{ij} \\ \bar{c}_{ij} & \bar{d}_{ij} \end{bmatrix}.$$

Because $\bar{a}_{ij}^2 + \bar{b}_{ij}^2 = 1$ and $\bar{c}_{ij}^2 + \bar{d}_{ij}^2 = 1$, then the elements of vectors $[\bar{a}_{ij}\ \bar{b}_{ij}]$, $[\bar{c}_{ij}\ \bar{d}_{ij}]$ can be coded as sines and cosines of two angles. Hence, the operator $\bar{A}_{ij}$ is equivalent to the operator (1c), i.e. $\bar{A}_{ij} \equiv B_{ij}$. Proving the same equivalence in the opposite direction is immediate. Based on the presented reasoning we find that each FPBNT with an additional scaling matrix D is equivalent to a certain biorthogonal transformation with the same structure (the first stage must be orthonormal) but exact to the form of operators and the values of coefficients in the second stage.

Definition of inverse transformation

On the basis of discussed in the previous section equivalence of biorthogonal (see [5]) and FPBNT structures we can easily formulate the inverse FPBNT transform. Let then $V = U_1 U_0$ be the forward FPBNT transformation. The inverse transform V^{-1} is constructed as:

1. the inverse matrix to the matrix U_1 of the second stage can be created on the basis of the definition of an inverse operator to B_{ij} (see [5]), i.e.:

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

where we assume: $a_{ij} = \cos(\alpha_{ij})$, $b_{ij} = \sin(\alpha_{ij})$, $c_{ij} = -\sin(\beta_{ij})$ and $d_{ij} = \cos(\beta_{ij})$. Hence:

$$A_{ij}^{-1} = \frac{1}{\cos(\alpha_{ij} - \beta_{ij})} \begin{bmatrix} \cos(\beta_{ij}) & -\sin(\alpha_{ij}) \\ \sin(\beta_{ij}) & \cos(\alpha_{ij}) \end{bmatrix}.$$

The operator A_{ij}^{-1} exists if only $\alpha_{ij} - \beta_{ij} \neq \frac{\pi}{2} + k\pi$, where $k \in \mathcal{Z}$. Knowing the form of A_{ij}^{-1} we can construct the matrix inverse to U_1 simply by executing the steps in reverse order and by replacing the operators A_{ij} with inverse operators A_{ij}^{-1},

2. the matrix inverse to the matrix of the first stage U_0 is its transpose U_0^T.

Experimental studies

In order to verify the practical effectiveness of the proposed FPBNT transforms a number of studies in the tasks of data classification were executed operating on the data set "Wine Data Set" available at [8]. The data set contains 178 vectors representing the attributes of three different species of wine. In this experiment we selected arbitrarily 8 from the set of 13 attributes in the form of: alcohol, malic acid content, ash density, alcalinity of ash, the level of magnesium content, phenols, flavonoids and proline content. Then we selected also in an arbitrary manner the training set consisting of 60 eight-element vectors, 20 vectors for each type of wine, while the remaining vectors served as the test set. All input vectors were shifted in the direction of the origin by the vector of mean values of attributes calculated only for the training set, and finally all vectors were normalized to the unit length.

The FPBNT with the number of $N = 8$ points and structure presented in Fig. 4 was applied to input data clustering by adapting transform coefficients to the training data set (note that the first stage U_0 is composed of orthonormal operators of rotation by the angle of 45° and is not subject to adaptation). In the process of FPBNT adaptation we utilized an algorithm inspired by the Kohonen neural network training algorithm, where the base vector that is closest in the sense of dot product to the input vector was modified. It should be noted that in the case of a fast structure modification of a single vector entails a change of the forms of the remaining base vectors.

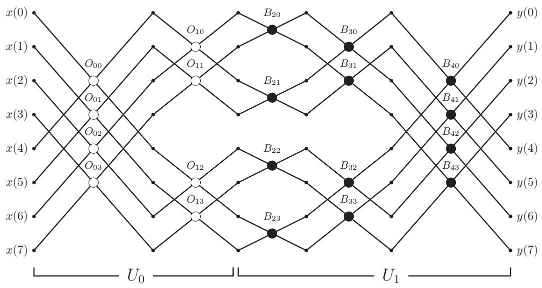

Fig. 4. The structure of FPBNT used in the task of data classification

After finishing the adaptation process the groups of representative vectors for each class of wines were determined and the classification process was carried out operating on the test set. We used for this purpose one-nearest neighbor classifier with a dot product as a distance measure. The results obtained with FPBNT in 1000 trials (see Table 1) in the form of mean classification accuracy (CA) and standard deviation (SD) were compared to the results obtained for the Kohonen neural network (KNN) with eight outputs and for $N = 8$ point fast parametrized orthonormal transform (FPT) with analogous structure (for definition of FPT see [2]).

Table 1. The classification results obtained in 1000 trails for FPBNT, the Kohonen neural network (KNN) and fast parametrized orthonormal transform (FPT) presented in the form of mean value of classification accuracy (CA) and its standard deviation (SD)

KNN		FPT		FPBNT	
CA [%]	SD [%]	CA [%]	SD [%]	CA [%]	SD [%]
88	1	67	10	85	4

An analysis of experimental results indicates the comparative behavior of the KNN and FPBNT. The worse at about 3% accuracy of classification, and higher standard deviation are a consequence of the limited capacity of the fast structure. However, this fact is compensated by lower computational requirements. In the considered case KNN and FPBNT required in order to calculate input vector's image the number of: 64 multiplications and 56 additions for KNN, but only 48 multiplications and 40 additions for FPBNT. With FPT we obtained significantly worse classification accuracy. This is due to strong restrictions on the perpendicularity of base vectors and such imperfection eliminates that transform from its practical applications to classification of data.

Conclusions

In this paper a scheme for constructing fast parametrized biorthogonal transforms with normalized basis vectors is proposed. The distinguishing feature of the proposed class of transforms is the unit length of base vectors, while the angles between vectors can take any values limited only by the fast computational structure.

The results obtained in practical task of data classification show comparable efficiency of the proposed transforms to the efficiency of Kohonen neural network, while keeping the fast computational scheme. This encourages the practical application of the proposed transforms especially on hardware platforms with poor computational performance, e.g. on mobile platforms.

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 Sig. Proc., 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] D. Puchala, M. Yatsymirskyy, *Fast parametrized biorthogonal transforms*, Electrical Review, vol. 4, pp. 123-125, 2012.

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

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

[8] A. Frank, A. Asuncion, *UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]*, Irvine, CA: University of California, School of Information and Computer Science, 2010.

Authors: Ph.D. Dariusz Puchala
Full 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