Face recognition using two-dimensional CCA and PLS

Georgy Kukharev

Department of Computer Software Environment,
Saint Petersburg Electrotechnical University 'LETI',
St. Petersburg, Russia
E-mail: kuga41@mail.ru

Andrzej Tujaka and Paweł Forczmański*

Faculty of Computer Science and Information Technologies,
West Pomeranian University of Technology, Szczecin,
Żołnierska 49, 71-210 Szczecin, Poland
E-mail: atujaka@neostrada.pl
E-mail: pforczmanski@wi.zut.edu.pl
*Corresponding author

Abstract: This paper presents the implementation of the method of two-dimensional Canonical Correlation Analysis (CCA) and two-dimensional Partial Least Squares (PLS) applied to image matching. Both methods are based on representing the image as the sets of its rows and columns and implementation of CCA using these sets (hence we named the methods as CCArc and PLSrc). CCArc and PLSrc feature simple implementation and lesser complexity than other known approaches. In applications to biometrics, CCArc and PLSrc are suitable to solving the problems when dimension of images (dimension of feature space) is greater than the number of images, i.e., Small Sample Size (SSS) problem. This paper demonstrates high efficiency of CCArc and PLSrc for a number of computer experiments, using benchmark image databases.

Keywords: CCArc; two-dimensional canonical correlation analysis; PLSrc; two-dimensional partial least squares; feature space dimensionality reduction; face matching.

Reference to this paper should be made as follows: Kukharev, G., Tujaka, A. and Forczmański, P. (2011) 'Face recognition using two-dimensional CCA and PLS', *Int. J. Biometrics*, Vol. 3, No. 4, pp.300–321.

Biographical notes: Georgy Kukharev received his PhD in 1979 and his Doctor of Technical Science in 1986. From 2006, he is a Full Professor. Now, he is a Professor at the Faculty of Computer Science and Information Technology West Pomeranian University of Technology (Szczecin, Poland) and a Professor in the Department of Computer Software Environment in St. Petersburg State Electrotechnical University (LETI), Russia. He is the author of 10 monographs, over 100 scientific papers in the following areas: computer architecture of signal processing, image processing and pattern recognition. His interests are in the areas of biometrics, including face detection, recognition and retrieval.

Andrzej Tujaka received his MSc in 1972 and PhD in 2001 from Electric Department and Faculty of Computer Science and Information Technology, respectively, of the Technical University of Szczecin (at present West Pomeranian University of Technology, Szczecin). His scientific interests include pattern recognition, especially methods of data dimensionality reduction in, e.g., biometrical applications, content-based image retrieval and visual surveillance He is the author of over 70 scientific articles, chapters in books, conference papers and technical reports.

Paweł Forczmański received his MSc in 1998 and PhD in 2002 from the Faculty of Computer Science and Information Technology, Technical University of Szczecin (at present West Pomeranian University of Technology, Szczecin). His scientific interests include digital image processing, recognition and analysis, methods of dimensionality reduction and image classification in biometrical applications, content-based image retrieval and visual surveillance. He is the author of over 50 scientific articles, chapters in books and conference papers.

1 Introduction

In recent years, Canonical Correlation Analysis (CCA) and Partial Least Squares (PLS) arouse the growing interest of experts in biometric recognition technologies, as a method allowing relating two or more blocks of observations on the same set of objects, each block describing different aspects of their appearance. In this framework, CCA and PLS transform information about set of objects from the space of originally observed features into the low-dimensional space of output variates, called in the case of CCA the Space of Canonical Variates (SCV) and in the case of PLS the Space of Components (SC), and all further processing is implemented in this space. Features in this space describe common information contained in two (or more) sets of different features, identified on a base of correlation and covariance between linear combinations of original features, making possible disclosure, investigation and modelling the latent associations between these sets.

Originally, the CCA method has been derived in 1936 by Hotelling for description of relations between two one-dimensional sequences of data (Hotelling, 1936). The PLS method has been proposed by Wold (1966) in the mid-1960s. Fast development of computer techniques – increasing memory capacity and processing speed, wide use of applied software packages for digital image processing (e.g., 'MATLAB', 'LabView' and 'Statistics'), application of mathematical modelling – has promoted application of these methods in processing the multidimensional data. Such data includes face images and images of, e.g., human gaits and hand gestures.

However, applications of CCA and PLS in image processing (Donner et al., 2006; Yi et al., 2007; Shan et al., 2008; Szaber and Kamenskaya, 2008; Alın et al., 2009) are usually connected with one-dimensional (vectorised) representation of original feature space. This means, for example, in case of image with $M \times N$ pixels, that dimension of representation will be equal to MN. For rather small images, where $M = 100$ and $N = 100$, Dimensionality of Feature Space (DFS) is equal to 10000 already. In practice of biometrics, according to the standard (www.icao.int/mrtd/download/technical.cfm), $\min\{M, N\} >> 100$, giving a value of DFS much more than 10,000. Such high

dimensionality makes direct application of CCA and PLS for image processing and biometrics impossible. In practice, it is the reason for using dimensionality reduction procedures, before using the CCA or PLS. Usually, it is implemented in two consecutive steps: first, reduction of input images size, then transformation of such images into vectors (Wegelin, 2000). Furthermore, the dimensionality reduction can be based on, e.g., Principal Component Analysis (Donner et al., 2006; Yi et al., 2007; Shan et al., 2008) or extraction of spectral features from original images (Szaber and Kamenskaya, 2008).

The aim of this paper is a comprehensive presentation of full algorithms for realising CCA and PLS and their novel implementations, specifically for two-dimensional input data (images). These methods are based on treating the image as a set of rows and a set of columns, and implementation of CCA and PLS based on these sets (hence the methods names are CCArc and PLSrc by analogy with Kukharev and Kuzminski (2003)).

The paper consists of the following sections: After Introduction showing some aspect of solving the face image recognition problem with the aid of CCA and PLS methods, we review shortly the history of two-dimensional extensions of standard methods used in image recognition and retrieval, like PCA, LDA, CCA and PLS. The next section briefly presents the standard CCA and PLS calculation structure; Section 4 presents structure of CCA and PLS methods extended on 2D data, called CCArc and PLSrc; Section 5 proposes implementation of CCArc and PLSrc methods and experimental results. Section 6 concludes the paper with some conclusions drawn from theoretical and experimental investigation of two-dimensional CCA and PLS methods used in face recognition.

2 An evolution of two-dimensional methods used in face image recognition problems

For the first time, the idea of treating the image as collection of its rows and columns was published in Tsapatsoulis et al. (1998), presenting vector-based approximation of PCA, in application to face recognition, called vector-based approximation of Karhunen-Loeve Transform (VKLT). The VKLT eliminates the large memory demands and the singularity of the covariance matrix, which are main disadvantages of the KLT. Authors claimed the VKLT is also performing better than ‘eigenface’ method under variation in illumination, translation and facial expression. Treating the image as collection of its rows and columns offers a framework greatly simplifying calculations needed in applications involving image data. Besides, it can be used in virtually all methods based on second-order statistics (i.e., scatter matrices) like PCA, LDA, CCA and PLS, unifying the required software.

Inspired by Tsapatsoulis et al. (1998), we have presented in Kukharev and Forczmanski (2001) a precise matrix form of this method, named 2DPCA (in relation to Yang et al. (2004)), specifying the size of involved matrices corresponding to face images. Complete algorithm of two-dimensional PCA has been presented taking into account the structure and parameters of face images database, specifying {K, L, Q-L, $M \times N$}, where K is a number of classes, L is a number of training images in each class, Q-L is a number of test images and $M \times N$ is a size of image matrix. We have presented an application to face recognition, combining 2DPCA with 1D LDA as well as to

recognition of satellite and aerial images. Methods presented earlier were implemented and tested in the MATLAB environment.

In the book Kukharev and Kuzminski (2003), we have separated the stage of 2DPCA from the two-dimensional KLT. We have also presented a graphical illustration of these stages of feature space dimensionality reduction, leading from input image of size $M \times N$ to image of size $p \times p$, where $p << \min\{M, N\}$. The whole procedure has been analysed and named as 2DPCA/2DKLT. We also presented some new applications of this procedure, as well-discussed recognition rate as a function of class number (K), number of images in each class (L), noise level and image scale. Robustness of procedure against the above-mentioned disturbances has been confirmed on benchmark data sets. Kukharev and Kuzminski (2003) describe the procedures implementing 1DPCA (local and global variant) and 2DPCA/2DKLT in Face Recognition Systems (FaReS). Because in the method 2DPCA/2DKLT row-row and column-column covariance matrices of images are used, we called this method PCArc/KLTrc. Basic feature of such algorithm is direct diagonalisation of both covariance matrices. This procedure has been implemented in a program package named FaReS-Mod, used for modelling the FaReS. In the paper (Kukharev and Forczmański, 2004), we have shown application of PCArc/KLTrc + 1DLDA for face recognition and described a scheme for specifying the algorithms.

Ye et al. (2005) and Ye (2005) presented GLRAM method, based on iterative 2DLDA procedure. Method does not use direct diagonalisation of the total scatter matrix.

Kukharev and Forczmanski (2005) have presented the algorithm 2DLDA/2DKLT using two covariance matrices (one for columns and one for rows of image matrix) for the first time. The presented algorithm explicitly took into account the structure and parameters of face images database. It has shown a graphical interpretation of this 2DKLT procedure of feature space dimensionality reduction, transforming the input image of size $M \times N$ into small image of size $s \times s$, where $s << \min\{M, N\}$. Jing et al. (2006) presented the algorithm named 2DFisherface, using two iterations, the first one implementing PCA along one dimension of input image, and the second one implementing LDA along the other dimension of the image obtained in the first iteration. Generally, the presentation is unclear, because in the figure illustrating 2DFisherface procedure, the block describing the first stage should implement 2DPCA, but formulas refer to LDA method applied to the columns of data matrix. The second stage shown in this figure is described as 2DLDA, but formulas refer to LDA procedure applied to the rows of matrix obtained in the first stage.

Kukharev and Forczmanski (2007) describe dimensionality reduction of face images, implemented by 2DKLT for two variants: 2DPCA and 2DLDA, and applied for face image database: FERET, ORL, BioID, Face94, Face95, mainly for the case when training set has more than one image. This paper contains also an evaluation of numerical complexity for method 2DLDA/2DKLT and the discussion of parameters FAR and FRR for described FaReS. It also demonstrates the influence of different kinds of deformations, such as folding down, tearing, blurred eyes, noise added to face image and on-face recognition effectiveness. Because of this fact, we have omitted description of the mentioned problems in the current paper.

Generally, the most distinctive feature of the two-dimensional methods is the calculation of covariance matrices directly using original image matrices. Existing 2D techniques can be categorised as follows (Nhat and Lee, 2007): one-sided low-rank approximation, using either column–column or row–row covariance matrices, and

two-sided low-rank approximation, using both column–column and row–row covariance matrices. Another categorisation is: iterative and non-iterative method. Methods presented in this paper belong to the two-sided non-iterative category.

3 Standard CCA and PLS calculation structure

We have in our disposal two sets of data, called X and Y, each consisting of K vectors of dimensionality $DIM \times 1$. The aim of both methods is finding projection matrices, transforming input data into the space of output variates (Hotelling, 1936; Wold, 1966; Alın et al., 2009):

$$U = W_x^T X, \quad V = W_y^T Y, \quad \text{so that} \quad \|U - V\| \to \min, \tag{1}$$

where U and V are output variates.

Canonical Correlation Analysis

In case of CCA, the output consists of two matrices U and V containing in columns canonical components with maximal correlation, equal the square root of eigenvalues of two simultaneous eigenproblems:

$$\begin{cases} (C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx}) w_x = \lambda_x w_x \\ (C_{yy}^{-1} C_{yx} C_{xx}^{-1} C_{xy}) w_y = \lambda_y w_y \end{cases}, \tag{2}$$

where C_{xx}, C_{yy}, C_{xy}, C_{yx} – covariance matrices of size DIM; λ appropriate eigenvalues; w_x and w_y eigenvectors (columns of the matrices W_x and W_y); $\lambda_x = 1/\lambda_y$.

Partial Least Squares

The aim of PLS is to find two projection matrices (1) and transforming input data into the space of new variates, i.e., the columns of matrices U and V (called the components, factors, or latent vectors), having maximal covariance. There are several different PLS techniques to extract latent vectors, and each of them gives rise to a variant of PLS. We seek such weight vectors w_x and w_y, for which PLS components have maximal covariance:

$$\begin{cases} (C_{xy} C_{yx}) w_x = \lambda_x w_x \\ (C_{yx} C_{xy}) w_y = \lambda_y w_y \end{cases}, \tag{3}$$

where all vectors and matrices are as defined in equation (2).

Both CCA and PLS end with selection of parameter d, specifying the reduction of input feature space dimensionality (where $d \le \min\{DIM, K\}$) and verification of the quality of mutual correspondence of variates in the ‘reduced’ features space. In image processing and biometric applications, the input information for CCA or PLS can be taken from features of input images (e.g., spectral components in Fourier base (Kukharev and Kuzminski, 2003) or basis of eigentransformation (Szaber and Kamenskaya, 2008)), anthropometric points of face, or elastic model of the face (Donner et al., 2006).

The possibility to solve the problem (2) and (3) and their numerical complexity depends on dimensionality of input feature vectors DIM, collected in data matrices X and

Y. Since the CCA and PLS analysis is based on the vector model of the input data, which in case of images leads to big sizes of involved covariance matrices, usually much bigger than the number of collected images, the preliminary dimensionality reductions procedure of the input feature space must be used (Donner et al., 2006; Yi et al., 2007; Shan et al., 2008; Szaber and Kamenskaya, 2008). Besides, reshaping image data into vectors might break their spatial structure and lost of important information.

This motivated the search for solutions allowing direct usage of image data as input, without reshaping them into vectors.

For instance, the papers Ho and Seungjin (2007) and Cai-rong et al. (2007) presented some variants of two-dimensional implementation of CCA in applications to face image processing, and paper Yang et al. (2008) presented a two-dimensional version of PLS used in face recognition. However, in Ho and Seungjin (2007) the solution is based on preliminary down-sampling (reducing the size of input images to 50×50 pixels), alignment of face according to the eyes line, and two consecutive iterative procedures for calculation of eigenvector matrices, requiring selection of its termination condition. Also in Cai-rong et al. (2007) the preliminary down-sampling of input images is used, reducing their size to 28×23 pixels, while the CCA procedure is implemented in one direction of the reduced images only – namely in row direction. Similar approach also presents the work (Yang et al., 2008).

Another solution to CCA method, using input images in a direct way, is presented in Kukharev and Kamenskaya (2009).

4 Structure of CCArc and PLSrc methods

CCArc and PLSrc methods, unlike approaches presented in Ho and Seungjin (2007) and Cai-rong et al. (2007), do not require preliminary down-sampling, are non-iterative and operate along both directions, i.e., along the rows and columns of input image, making it different from Cai-rong et al. (2007).

Let us consider two sets, consisting of K pairs of images of size $M \times N$ pixels:

$$X = [X^{(1)} X^{(2)} \ldots X^{(K)}] \quad \text{and} \quad Y = [Y^{(1)} Y^{(2)} \ldots Y^{(K)}], \tag{4}$$

where $X^{(k)}$, $Y^{(k)}$ $\forall k \in K$ – pairs of input images.

Methods of CCArc and PLSrc are described in the following steps:

Average images $\bar{X}$ and $\bar{Y}$ are calculated for each input data set:

$$\bar{X} = \frac{1}{K} \sum_{k=1}^{K} X^{(k)} \quad \text{and} \quad \bar{Y} = \frac{1}{K} \sum_{k=1}^{K} Y^{(k)}; \tag{5}$$

where $\bar{X}, \bar{Y}$ are matrices of size $M \times N$.

Let the matrices $\boldsymbol{X}1$ and $\boldsymbol{Y}1$, as well $\boldsymbol{X}2$ and $\boldsymbol{Y}2$ contain the ‘rows’ and ‘columns’ of all centred input images:

$$\begin{aligned} &X1 = [\bar{X}^{(1)} \bar{X}^{(2)} \ldots \bar{X}^{(k)}], \quad Y1 = [\bar{Y}^{(1)} \bar{Y}^{(2)} \ldots \bar{Y}^{(K)}], \\ &X2 = [(\bar{X}^{(1)})^T \ \ldots \ (\bar{X}^{(k)})^T], \quad Y2 = [(\bar{Y}^{(1)})^T \ldots (\bar{Y}^{(K)})^T] \end{aligned} \tag{6}$$

where $\bar{X}^{(k)} = X^{(k)} - \bar{X}$ and $\bar{Y}^{(k)} = \bar{Y}^{(k)} - \bar{Y}$, $\forall k$, and the matrices $\boldsymbol{X}\mathbf{1}$ and $\boldsymbol{Y}1$ have the size $M \times NK$, and matrices $\boldsymbol{X}2$ and $\boldsymbol{Y}2$ are of size $N \times MK$.

In that case, the covariance matrices of data (6) can be presented as follows:

$$\begin{aligned} &C_{xx}^{(r)} = X1(X1)^T; \quad C_{yy}^{(r)} = Y1(Y1)^T; \quad C_{xy}^{(r)} = X1(Y1)^T \\ &C_{xx}^{(c)} = X2(X2)^T; \quad C_{yy}^{(c)} = Y2(Y2)^T; \quad C_{xy}^{(c)} = X2(Y2)^T. \end{aligned} \tag{7}$$

For equation (7), we evaluate total matrices, for rows (r) and for columns (c) separately.

For CCArc, it can be determined in the following manner:

$$\begin{aligned} S^{(\text{total1},r)} &= [C_{xx}^{(r)}]^{-1} C_{xy}^{(r)} [C_{yy}^{(r)}]^{-1} C_{yx}^{(r)}; \\ S^{(\text{total2},r)} &= [C_{yy}^{(r)}]^{-1} C_{yx}^{(r)} [C_{xx}^{(r)}]^{-1} C_{xy}^{(r)}; \\ S^{(\text{total1},c)} &= [C_{xx}^{(c)}]^{-1} C_{xy}^{(c)} [C_{yy}^{(c)}]^{-1} C_{yx}^{(c)}; \\ S^{(\text{total2},c)} &= [C_{yy}^{(c)}]^{-1} C_{yx}^{(c)} [C_{xx}^{(c)}]^{-1} C_{xy}^{(c)}, \end{aligned} \tag{8}$$

where $C_{yx}^{(r)} = [C_{xy}^{(r)}]^T$ and $C_{yx}^{(c)} = [C_{xy}^{(c)}]^T$.

For PLSrc, it can be determined in the following manner:

$$\begin{aligned} &S^{(\text{total1},r)} = C_{xy}^{(r)} [C_{xy}^{(r)}]^T, \quad S^{(\text{total2},r)} = [C_{xy}^{(r)}]^T C_{xy}^{(r)}; \\ &S^{(\text{total1},c)} = C_{xy}^{(c)} [C_{xy}^{(c)}]^T; \quad S^{(\text{total2},c)} = [C_{xy}^{(c)}]^T C_{xy}^{(c)}. \end{aligned} \tag{9}$$

Matrices (8) and (9) are diagonalised, and thus eigenvalues are determined with corresponding eigenvectors, contained in matrices $\mathbf{\Lambda}$ and $\boldsymbol{W}$:

$$\begin{aligned} &\text{For PLS: } \begin{cases} S^{(\text{total},r)} W_{x_1} = W_{x_1} \Lambda_x^{(r)}; \\ S^{(\text{total},r)} W_{y_1} = W_{y_1} \Lambda_y^{(r)}; \end{cases} \\ &\text{For CCA: } \begin{cases} S^{(\text{total1},r)} W_{x_1} = W_{x_1} \Lambda_x^{(r)}; \\ S^{(\text{total2},r)} W_{y_1} = W_{y_1} \Lambda_y^{(r)}; \end{cases} \quad \text{and} \quad \begin{cases} S^{(\text{total1},c)} W_{x_2} = W_{x_2} \Lambda_x^{(c)}; \\ S^{(\text{total2},c)} W_{y_2} = W_{y_2} \Lambda_y^{(c)}. \end{cases} \end{aligned} \tag{10}$$

Note that matrices with superscripts (r) and (c) in formulas (5)–(10) have order (rank) M and N, respectively. This ends the first stage of PLSrc and CCArc.

Second stage – transformation of input data into new feature space and dimensionality reduction. This is implemented as two-dimensional Karhunen-Loeve transformation (2D KLT), which can be presented in the following form:

$$U^{(k)} = W_{x_1}^T \overline{X}^{(k)} W_{x_2} \quad \text{and} \quad V^{(k)} = W_{y_1}^T \overline{Y}^{(k)} W_{y_2}, \quad \forall k \in K, \tag{11}$$

where $U^{(k)}$, $V^{(k)}$ are matrices containing data representations in new spaces.

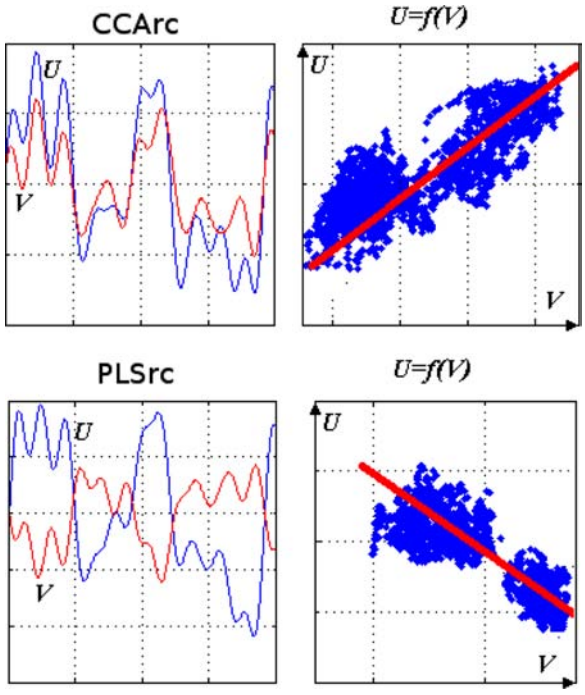
The size of matrices in equation (11) corresponds to the original size of input images, making those matrices of output variates, defined as:

$$U = [U^{(1)} U^{(2)} \ldots U^{(K)}] \quad \text{and} \quad V = [V^{(1)} V^{(2)} \ldots V^{(K)}] \tag{12}$$

are of size $M \times NK$.

Graphical representation of relation between output variates in the two-dimensional plane for CCArc and PLSrc are presented in Figure 1.

Figure 1 The graphical representation of output variates (see online version for colours)



As it can be seen, variates *U* and *V* in the CCArc are associated with each other, resulting in maximal correlation, which is the primary aim of this method. At the same time in PLSrc, variates *U* and *V* are similar, but in 'opposite phase'. Such superposition leads to increased dispersion between variates *U* and *V*, which is the goal of PLSrc.

From the other side, in the new space, variates *U* and *V* are linearly dependent, allowing building a regressional model of variables *U* and *V*. Such models are used to form mutual mapping between patterns. The presented variates are calculated for images taken from Equinox database described in the further part of the paper.

Now, we can determine the 'principal' components. These will be the first *d* biggest eigenvalues in equation (10). Sorting the eigenvalues in descending order, then for arbitrary $d \geq 1$, the following conditions are always met:

$$\sum_{i=1}^{d_1} (\lambda_i^{(r)})^2 >> \sum_{i=d_1+1}^{M} (\lambda_i^{(r)})^2 \quad \text{and} \quad \sum_{i=1}^{d_2} (\lambda_i^{(c)})^2 >> \sum_{i=d_2+1}^{N} (\lambda_i^{(c)})^2, \tag{13}$$

where $\lambda_i^{(r)}$ and $\lambda_i^{(c)}$ are eigenvalues; $d_1 << M$; $d_2 << N$ and $d_1 \neq d_2$ generally.

Lower bound of parameter *d* is selected experimentally, in verification stage of the CCArc and PLSrc – using the criteria equation (1).

To achieve necessary effect of DFS reduction, the procedure (11) should be modified so that in the projection the eigenvectors corresponding to *d* 'principal' variates and components are used.

To achieve this aim, from matrices $[W_{x_1}]^T$ and $[W_{y_1}]^T$ we select *d* rows, corresponding to *d* biggest eigenvalues, from which we form reduction matrices F_{x_1} and

F_{y_1}. From matrices W_{x_2} and W_{y_2}, we select d columns, corresponding to d biggest eigenvalues, and from them we form matrices F_{x_2} and F_{y_2}.

Now, for each centred input image, we apply a truncated 2D KLT, which can be presented in the following form:

$$\hat{U}^{(k)} = F_{x_1} \overline{X}^{(k)} F_{x_2}; \quad \hat{V}^{(k)} = F_{y_1} \overline{Y}^{(k)} F_{y_2}, \quad \forall k \in K. \tag{14}$$

The variates with a hat emphasise the difference of the above-mentioned result in respect to equation (11).

Size of resulted matrices in equation (14) is equal to $d \times d$ or $d_1 \times d_2$ in general case. Obtained matrices of output data are determined as:

$$\hat{U} = [\hat{U}^{(1)} \hat{U}^{(2)} \ldots \hat{U}^{(K)}] \quad \text{and} \quad \hat{V} = [\hat{V}^{(1)} \hat{V}^{(2)} \ldots \hat{V}^{(K)}] \tag{15}$$

and will have the size of $d \times dK$ or $d_1 \times d_2 K$.

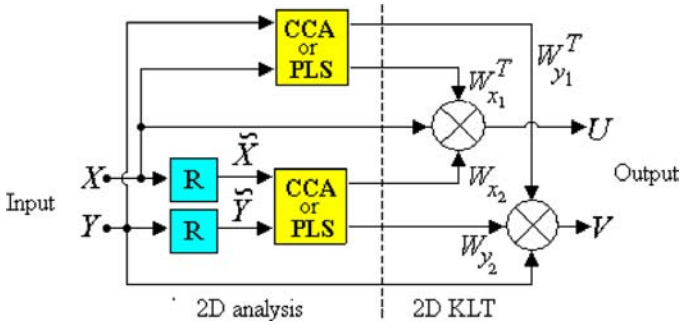
This ends the second stage and the whole procedure of CCArc and PLSrc.

5 Practical aspects of CCArc and PLSrc

5.1 Numerical implementation

Details of the two-dimensional CCA and PLS (within CCArc and PLSrc) can be observed in Figure 2. Boxes designated as '⊗' implement the projection procedures of input data into the space of canonical or new variates.

Figure 2 Simplified scheme of two-dimensional CCA and PLS procedures (see online version for colours)



For solving the eigenproblems, we use regularisation of matrix C_{xx}, C_{yy}, as well as total matrices, as described, e.g., in Sun et al. (2009).

It can be seen that CCArc and PLSrc do not require any preliminary down-sampling and they are applied along both directions – namely along rows and columns of input image.

5.2 Evaluation of numerical complexity of two-dimensional CCA and PLS

Note that for implementation of 2D KLT according to equation (11) we need $(NM^2 + MN^2)$ operations, and implementation of truncated 2D KLT according to

equation (15) needs ($Md^2 + MNd$) operations. Reduction of number of operations in the stage of 2D KLT can be approximately evaluated as $(M + N)/d$, since:

$$\frac{NM^2 + MN^2}{Md^2 + MNd} = \frac{MN(M+N)}{M(d+N)d} \approx \frac{M+N}{d}, \quad \text{if we assume } d << N. \tag{16}$$

Results of equation (16) determine the speed-up of calculations for each of the processed image. Using the parameter K – number of images in each of both sets, acceleration of calculation can be expressed as $2K(M + N)/d$ for all input data.

Matrices $\hat{U}^{(k)}, \hat{V}^{(k)}$ in equation (15) contain d^2 elements, i.e., MN/d^2 times less than the number of elements in input image. It means that reduction degree is determined by the ratio MN/d^2. For example, for $M = 112$, $N = 92$ (size of images in ORL database) and $d = 10$, dimensionality reduction will amount to more than 100 times (!) for each image.

Note also that the biggest size of covariance matrix in equations (8) and (9) is $DIM = \max\{M, N\}$, which allows us to solve the eigenvalue problem in practice, with high stability of such solution even for the size of images defined in standard (www.icao.int/mrtd/download/technical.cfm). Usually, the size of these matrices is connected with the size of original images, often equal to $MN >> DIM = \max\{M, N\}$.

Finally, the problem of SSS (Kukharev and Forczmański, 2004, 2007; Ho and Seungjin, 2007), when $DIM > K$, in CCArc or PLSrc is also eliminated, since instead of one image of $M \times N$ pixels we actually have N images of size $M \times 1$ and M images of size $N \times 1$. For such representation of input data, we always satisfy the condition: $DIM = \max\{M, N\} < (M + N)$, which was stressed many times in Kukharev and Forczmański (2004, 2007), Kukharev and Kuzminski (2003) and Kukharev and Kamenskaya (2009).

5.3 Experiments

Experiments were carried out in the environment of MATLAB software package using different architectures of FaReS, which has already been investigated in Szaber and Kamenskaya (2008), Kukharev and Kuzminski (2003). The aim of conducted experiments was the dimension reduction of input data using the CCArc and PLSrc methods, and the face recognition task in these reduced spaces. Experiments were executed for two variants of databases.

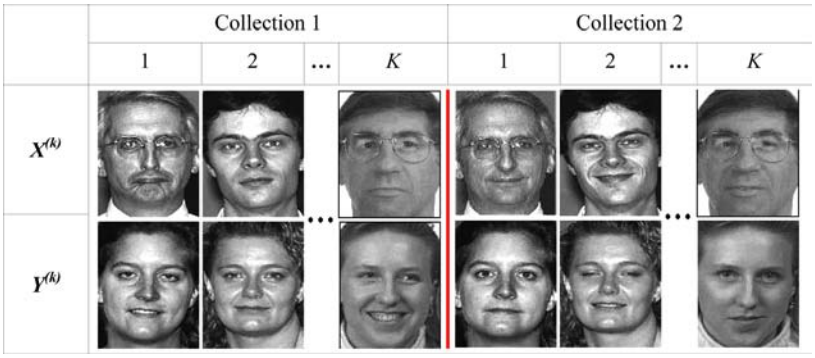
For the first experiment, we used developed ‘Family Album’ database (DBFA), comprising portraits of men and women, extracted from FERET (Philips et al., 1998) database. Furthermore, these pairs of images are called ‘Husband’ and ‘Wife’ pairs.

For the second experiment, we have prepared database of face images (DB VIS/Th), composed of images from Equinox database (http://www.equinoxsensors.com/products/HID.html), containing sets of image pairs belonging to two categories – images in visible and near infrared light.

5.3.1 Experiments on Family Album (DBFA) database

The DBFA database is partitioned into two collections, having the structure presented in Figure 3, where $X^{(k)}$ and $Y^{(k)}$ – are defined as pair of images of ‘Husband’ and ‘Wife’, for $\forall k \in K$. In this case, $K = 100$, and each image size is 224×184 pixels, and remains constant during all experiments.

Figure 3 The structure of 'Family Album' database (see online version for colours)



Each pair of images in Collection 1 has a corresponding pair in Collection 2, differing in face expression, and slight variations in lighting, head size and orientation. Images from Collection 1 are used for training of FaRES, and images from Collection 2 are used for testing (and vice versa). In this scheme, we can, e.g., using the ordering procedures for one set only (e.g., ***Y***), obtain up to 10,000 various variants with 100 classes of 'Husband' and 'Wife'. This variant of Cross-Validation (CV) has been used in described experiments.

The aim of experiment is to solve the task of face image retrieval of class 'Wife' based on the image of 'Husband' and vice versa. Formally, it is the task of image retrieval from set Y based on given image from set X (and vice versa). Note that face image pair $X^{(k)}$, $Y^{(k)}$ for each k can be treated as a separate class.

This is the difficult problem, because we cannot expect large correlation of images from the same class, formed by 'Wife' and 'Husband' images. But in the new feature space this correlation is enhanced, making a solution of given task easier.

Using the notation analogous to used in Szaber and Kamenskaya (2008), we can describe the models of conducted experiments as:

$$\text{DBFA}(100/(1(X) + 1(Y))/1//\text{CV})\{\text{CCArc}{:}224 \times 184 \rightarrow (d \times d)/\text{MDC}/\text{L}_1/rank = 1\};$$

$$\text{DBFA}(100/(1(X) + 1(Y))/1//\text{CV})\{\text{PLSrc}{:}224 \times 184 \rightarrow (d \times d)/\text{MDC}/\text{L}_1/rank = 1\},$$

where the following conditions are specified:

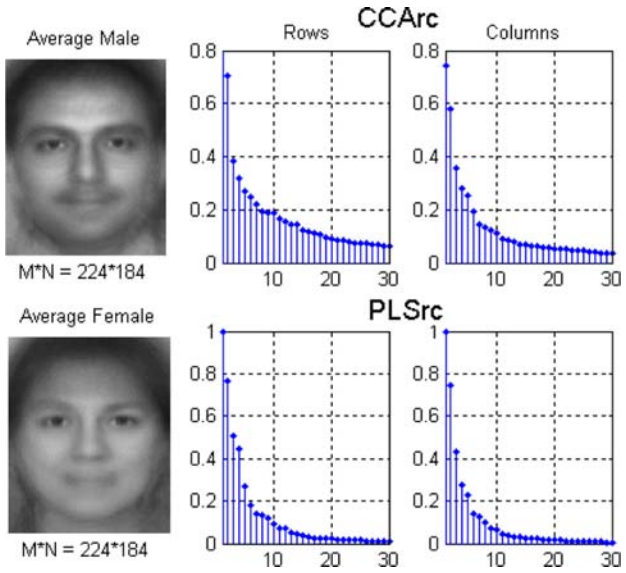
- structure of input data
- used CV procedure
- dimension reduction method CCArc (or PLSrc)
- dimension of input and reduced spaces (here $d = 5 \div 40$ with step of 5)
- classification method (here Minimum Distance Classifier – MDC), used metric (L_1), and the number (*rank)* of retrieved images.

In retrieval task, the database of FaReS contained 100 pairs of training images X and Y, and as test images were used 100 images (one image X or Y for each of 100 classes). The results obtained in each experiment were assessed as the ratio of correctly classified test images to the number of test images used, i.e., 100 test images. In classification procedure, we used the metric L_1 for measuring the distance of test image to average

image in the class (MDC/L_1). The result is regarded as correct if in the first place (*rank* = 1) in the group of retrieval result is the image from the same class as test image.

Figure 4 presents average images for the whole training set, and descending ordered set of the square root of eigenvalues (first 30 values) $\sqrt{\lambda^{(r)}}$ and $\sqrt{\lambda^{(c)}}$, evaluated according to equations (8) and (10).

Figure 4 Distribution of eigenvalues in CCArc and PLSrc methods (see online version for colours)



In this case, these values express the relationship between the corresponding columns (and rows) of the images 'Husband' and 'Wife' in the space of reduced dimensions (in the space of canonical correlations for the CCA method).

Table 1 presents the results of searching for optimal value of parameter *d*, giving the best verification and retrieval results of the image *X* based on image *Y* (and vice versa). These results are contained in the rows 'Verification' and the rows test '$Y \rightarrow X$' and '$X \rightarrow Y$', respectively. Table 1 shows that the best results are obtained with $d = 20$. In this case, the dimension reduction ratio of feature space is $(224 \times 184)/(20 \times 20) > 100$.

Table 1 Comparison of the results of retrieval for CCArc and PLSrc

Method	*DBFA 100/(1 + 1)/1*								
	d	*5*	*10*	*15*	*20*	*25*	*30*	*35*	*40*
CCArc	Verification %	77	96	*100*	*100*	*100*	*100*	*100*	*100*
	Test '$X \rightarrow Y$' %	51	69	76	78	73	72	72	70
	Test '$Y \rightarrow X$' %	46	79	86	85	83	80	78	74
PLSrc	Verification %	82	99	*100*	*100*	*100*	*100*	*100*	*100*
	Test '$X \rightarrow Y$' %	65	87	90	92	91	92	90	90
	Test '$Y \rightarrow X$' %	63	86	87	89	89	89	89	88

Table 2 shows the results of tests '$Y \rightarrow X$' and '$X \rightarrow Y$' for the value of $d = 20$ and ten variants of input data, obtained in CV procedure. In this procedure, we generated ten variants of 'Family Album', including 100 new pairs of training images and corresponding 100 pairs of testing images.

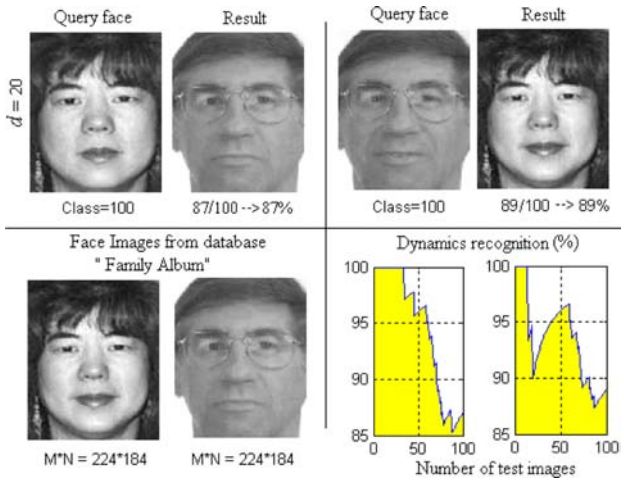
Table 2 Results of recognition for $d = 20$

Method	*Experiment*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Average*	
CCArc	Test 'X → Y' %	79	80	81	78	77	79	82	79	80	82	*79.7*	*80*
	Test 'Y → X' %	85	84	85	80	72	82	79	81	83	79	*81*	
PLSrc	Test 'X → Y' %	92	89	90	89	92	90	89	89	88	91	*89.9*	*89*
	Test 'Y → X' %	88	87	86	86	90	86	91	87	87	87	*87.5*	

We can see from Table 2 that the average value of mutual recognition (i.e., spouses) attained 80% for the method CCArc and 89% for the method PLSrc.

Figure 5 presents one of the results of spouse retrieval procedure in the 'Family Album'. Right bottom part of figure presents the dynamic of mutual recognition of spouses in running the tests.

Figure 5 Results obtained with spouse retrieval procedure (see online version for colours)



As results from conducted experiments, neither the PLSrc nor CCArc method obtained 100% of successes of retrieval of the image from *Y*, based on given image from the set *X* (and vice versa). This is the consequence of the fact that images contained in 'Family Album', forming the 'pair of spouses', from the definition belong to different persons and genders, and cannot form the general classes. However in the frame of multi-biometric tasks, being the generalisation of presented task, we have exactly such situation.

We can qualitatively assess the results presented earlier by comparing them with the results of other method. Hence, we can use the two-dimensional Discrete Fourier Transform method (2D DFT) (Kukharev and Kuzminski, 2003), instead of PLSrc and CCArc.

This method has been selected because in 2D DFT the transformation of input data into the new feature space (spectral space) is also conducted along the columns and rows of the images, using the procedures analogous to equation (14). 2D DFT has been implemented based on the following matrix form:

$$C = \text{abs}(F^{(c)} X^{(k)} F^{(r)}) \quad \forall k \in K, \tag{17}$$

where

$$F^{(c)} = \begin{bmatrix} \omega_M^0 & \omega_M^0 & \ldots & \omega_M^0 \\ \omega_M^0 & \omega_M^1 & \ldots & \omega_M^{M-1} \\ \omega_M^0 & \omega_M^2 & \ldots & \omega_M^{2(M-1)} \\ . & . & \ldots & . \\ \omega_M^0 & \omega_M^{(p-1)} & \ldots & \omega_M^{(p-1)(M-1)} \end{bmatrix} \quad \text{matrix size } p \times M;$$

$$F^{(r)} = \begin{bmatrix} \omega_N^0 & \omega_N^0 & \ldots & \omega_N^0 \\ \omega_N^0 & \omega_N^1 & \ldots & \omega_N^{p-1} \\ \omega_N^0 & \omega_N^2 & \ldots & \omega_N^{2(p-1)} \\ . & . & \ldots & . \\ \omega_N^0 & \omega_N^{(N-1)} & \ldots & \omega_N^{(N-1)(p-1)} \end{bmatrix} \quad \text{matrix size } N \times p;$$

$$\omega_N^L = \exp\left(-j\frac{2\pi}{N}L\right) = \cos\left(\frac{2\pi}{N}L\right) - j\sin\left(\frac{2\pi}{N}L\right), \quad \forall L = 0,1,\ldots,N;$$

$$\omega_M^L = \exp\left(-j\frac{2\pi}{M}L\right) = \cos\left(\frac{2\pi}{M}L\right) - j\sin\left(\frac{2\pi}{M}L\right), \quad \forall L = 0,1,\ldots,M.$$

Each transformation matrix (17) consists of two parts:

$$F^{(c)} = F_{p\times M}^{\text{real}} - jF_{p\times M}^{\text{imag}} \quad \text{and} \quad F^{(r)} = F_{N\times p}^{\text{real}} - jF_{N\times p}^{\text{imag}}, \tag{18}$$

where the part (real) and (imag) of the matrix (18) are determined by the components (cos) and (sin), respectively. Hence, the resulting matrix in equation (17) can also be written in the following form:

$$C = C_{p\times p}^{\text{real}} - jC_{p\times p}^{\text{imag}}. \tag{19}$$

This implies that the result in equation (17) in real number arithmetic shall have the following form:

$$C_{p\times p}^{\text{real}} = F_{p\times M}^{(\text{real})} X_{M\times N} F_{N\times p}^{(\text{real})} - F_{p\times M}^{(\text{imag})} X_{M\times N} F_{N\times p}^{(\text{imag})}; \tag{20}$$

$$C_{p\times p}^{\text{imag}} = F_{p\times M}^{(\text{real})} X_{M\times N} F_{N\times p}^{(\text{imag})} + F_{p\times M}^{(\text{imag})} X_{M\times N} F_{N\times p}^{(\text{real})}.$$

Finally, in correspondence with the problem of computing the module of the spectrum, based on equation (17), we have the real matrix $\hat{C}_{p\times p}$, presenting the final result:

$$\hat{C}_{p\times p} = \mathrm{abs}(C_{p\times p}) = \sqrt{(C^{\mathrm{real}}_{p\times p})^2 + (C^{\mathrm{imag}}_{p\times p})^2}. \tag{21}$$

Figure 6 presents the scheme of 2D DFT implementation, based on procedures (17)–(20). Two blocks in this figure correspond to real and imaginary matrices from equation (18). Sizes of blocks are determined by the parameters $\{M, N, p\}$ as shown in the figure.

To improve the results, more experiments were conducted with various number of family pairs, starting with $K = 2$. Hence, in accordance with adopted form of the FaReS system description, model of executed experiment is denoted as:

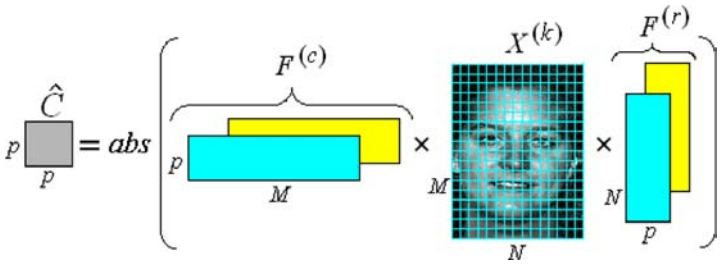
$$\mathrm{DBFA}(2 \div K/(1(X) + 1(Y))/1)\{\mathrm{2D\ DFT}{:}224 \times 184 \rightarrow (20 \times 20)/\mathrm{MDC}/\mathrm{L}_1/\mathit{rank} = 1\},$$

where

K: Number of family pairs (K = 2, 4, 6, 8, 10, 15, 20,…, 100)

p: Dimension of the new features space, specifically p = 20.

Figure 6 Scheme of 2D DFT implementation (see online version for colours)



The value of p = 20 has been selected intentionally, since to have exact representation of input image of face it is enough to use 20 spatial components of the DFT (Kukharev and Kuzminski, 2003). Besides, the value of p = 20 corresponds to the parameter d = 20, as used in the above-mentioned experiments with CCArc and PLSrc.

We have also conducted the analogous experiment using the CCArc and PLSrc methods, which models can be described as:

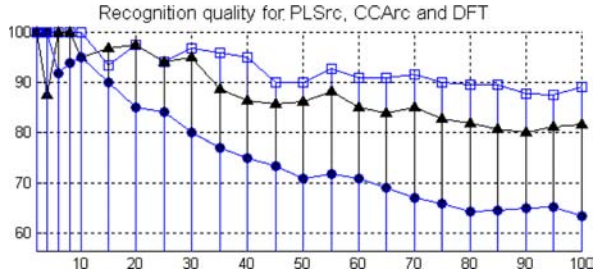
$$\mathrm{DBFA}(2 \div K/(1(X) + 1(Y))/1)\{\mathrm{CCArc}{:}224 \times 184 \rightarrow (20 \times 20)/\mathrm{MDC}/\mathrm{L}_1/\mathit{rank} = 1\};$$

$$\mathrm{DBFA}(2 \div K/(1(X) + 1(Y))/1)\{\mathrm{PLSrc}{:}224 \times 184 \rightarrow (20 \times 20)/\mathrm{MDC}/\mathrm{L}_1/\mathit{rank} = 1\}.$$

Figure 7 presents results of these three experiments. Here, the ordinate axis scores are the correct classification rate (in %), and abscissa axis defines the current values of parameter K – number of classes (family pairs) in the DB of FaReS. Symbols '□', 'Δ' and '•' denote the results for PLSrc, CCArc and DFT, respectively.

The results presented earlier show qualitatively that PLSrc and CCArc gave substantially better results than the discrete two-dimensional Fourier transform. This difference is clearly visible starting from the number of class $K \geq 15$. With increased number of classes in the database of FaReS, the difference in results obtained with PLSrc (CCArc) and 2D DFT is becoming more distinct.

Figure 7 Correct recognition rates of spouse retrieval vs. the number of family pairs, for three methods (see online version for colours)

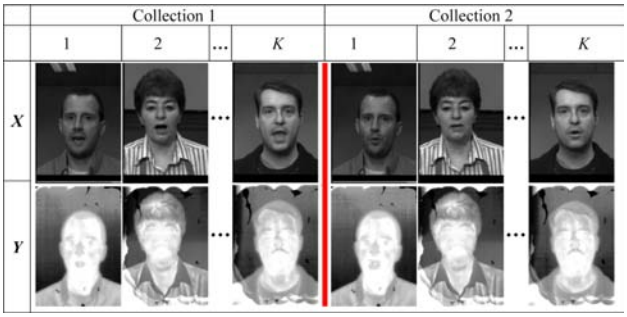


Finally, we would like to stress one more important fact. The presented results have shown that the methods CCArc and PLSrc work also in cases when the size of image (and the dimension of input features space) is greater than the number of images (the problem of SSS (Forczmanski and Kukharev, 2007)). In our case, these sizes were as follows: $DIM = 224 \times 184$ and $2 \le K \le 100$. It is noteworthy to stress that implementation of CCArc and PLSrc does not require meeting the condition $DIM < K$, making possible their wider applications in big size image processing problems, including ones conforming the requirements of the standard (http://www.equinoxsensors.com/products/HID.html). The results obtained in the above-mentioned experiments are in favour of such a claim. Additionally, it will be supported by the results of experiments described here.

5.3.2 Experiments on DBVIS/Th

The structure of a database of face images DBVIS/Th (originally Equinox database (http://www.equinoxsensors.com/products/HID.html)) is illustrated in Figure 8. The database consists of two collections, each having the structure (2).

Figure 8 Images forming DBVIS/Th, created from the Equinox database (see online version for colours)



Here, the images $X^{(k)}$ and $Y^{(k)}$ for $\forall k \in K$ are the frontal upright portraits taken in visible light (VIS) and corresponding portraits in near infrared light (Th). The number of persons is $K = 90$, each image has the size 320×240 pixels, and remains the same in our experiments. Note that the size of 320×240 pixels is the smallest size of human face allowed in biometric applications (see standard (www.icao.int/mrtd/download/technical.cfm)).

Each pair of images (pair VIS and Th) from Collection 1 corresponds to the pair of images in Collection 2, which differs in background luminance, face expression and small variations in size and orientation of head. Face images from Collection 1 are used for training the FaReS, whereas images from Collection 2 are used for testing (and vice versa). Besides, all images in database DBVIS/Th have different 'artefacts', created by specific conditions of taking the images for "Equinox IR Face Database".

Examples of such artefacts are shown in Figure 9, and in our experiment they will be not removed from original images.

Shown 'artefacts' can be described as follows:

Areas '1' and '2' – white areas, wide vertical and horizontal white and black stripes, in the images of category 'VIS';

- objects '3' – dark 'figure-like' areas and 'thermal cloud' in the images of category 'Th'
- changed colours '4' – differences in luminance and background in some areas of images category VIS and Th
- hidden information '5' – 'replacement' of visible part of face close to eyes by dark area in Th – images.

Note that in CCArc and PLSrc we evaluate the covariance matrices, which should have rank equal to the size of matrix. However, in the presence of artefacts in original images of types 1 and 2 (see Figure 9 and accompanying description) rank of covariance matrices will be smaller than their size, because in original images we have whole rows and columns filled with constant values, equal to the minimal (0) or maximal value (255). Rank of covariance matrices will not be the same, and will depend on the number of such rows and columns in individual images. This could worsen the quality of obtained result, and even lead to situations when covariance matrices cannot be used or making the solution of recognition problem impossible. We can avoid such problems, using some 'improvements' of original data – e.g., to add some noise to every image of category VIS. This approach has been used in our experiments, and because of this there is no need to remove artefacts of types 1 and 2 from the original images.

Figure 9 Examples of artefacts on images in face database DBVIS/Th (see online version for colours)

Taking into account the above-mentioned conditions, the aim of experiments is an attempt to solve the retrieval task of image from class 'VIS' based on the image from class 'Th' and vice versa. In a formal way, we will be solving the problem of retrieving the image from the set $\boldsymbol{Y}$ based on given image from the set X (and vice versa). Note that the pair of face images $X^{(k)}$, $Y^{(k)}$ for all k is treated here as separate classes.

Models of conducted experiments for CCArc and PLSrc, expressed in notations used earlier, are as follows:

$$\text{DBVIS/Th}(90(1(X + N^{(d)})+1(Y))/1)\{\text{CCArc:}320 \times 240 \rightarrow (d \times d)/\text{MDC}/L_1/rank = 1\};$$

$$\text{DBVIS/Th}(90(1(X + N^{(d)}) + 1(Y))/1)\{\text{PLSrc:}320 \times 240 \rightarrow (d \times d)/\text{MDC}/L_1/rank = 1\},$$

where '$X + N^{(d)}$' denotes the procedure of 'improvement' of original images of category VIS.

Generation of 'new' images $X^{(k)}$ is made according to the formula $X^{(k)} = X^{(k)} + N^{(k,d)}$, $\forall k \in$ K and for all of eight experiments, connected with changing the parameter d (here d = 5, 10, …, 40). In the last formula, size of matrix $N^{(k,d)}$ of random number is equal to the size of $X^{(k)}$.

Table 3 presents results of selecting the parameter d, ensuring the best result of verification and classification of images VIS ↔ Th: tests 'VIS → Th' and 'Th → VIS'.

Table 3 The results of recognition for different parameters d

Method	*DBVIS/Th(89/(1 + 1)/1*								
	Parameter d	*5*	*10*	*15*	*20*	*25*	*30*	*35*	*40*
CCArc	Verification %	68	98	100	*100*	100	100	100	100
	Test 'VIS → Th' %	50	85.6	90	*92.2*	92.2	90	90	88
	Test 'Th → VIS' %	30	93.3	95.6	*94*	95.6	95.6	95.6	94.4
	Average	*40*	*89.4*	*92.8*	*93.8*	*93.8*	*92.8*	*92.8*	*92.2*
PLSrc	Verification %	33	78	91	97	98	100	100	100
	Test 'VIS → Th' %	37	72	86	93	96	97	97	97
	Test 'Th → VIS' %	33	76	81	90	96	96	96	96
	Average	*35*	*74*	*83*	*92*	*96*	*96*	*96*	*96*

As we can see, the best result is obtained with CCArc and value of the parameter d = 20, and for PLSrc method with the parameter value d = 30.

Table 4 gives results of tests '$Y \rightarrow X$' and '$X \rightarrow Y$' for the value d = 20 (for CCArc) and d = 20, 30 (for PLSrc) and 10 variants of input data.

For obtaining the representative results, we used the procedure described earlier $\forall k \in K$ and in each ith experiment (i = 1, 2, …, 10), and averaging the results. Models of conducted experiments in this case are as follows:

$$\text{DBVIS/Th}(90(1(X + N) + 1(Y))/1)\{\text{CCArc:}320 \times 240 \rightarrow (d \times d)/\text{MDC}/L_1/\text{rank} = 1\};$$

$$\text{DBVIS/Th}(90(1(X + N) + 1(Y))/1)\{\text{PLSrc:}320 \times 240 \rightarrow (d \times d)/\text{MDC}/L_1/\text{rank} = 1\}.$$

Table 4 The results of recognition for optimal *d* values

Method	*Experiment*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Average*	
CCArc	Test '$X \to Y$' %	90	92	90	92	97	90	93	90	92	91	*91.7*	*93.7*
d = 20	Test '$Y \to X$' %	96	96	94	96	96	96	96	96	96	94	*95.6*	
PLSrc	Test '$X \to Y$' %	90	91	92	94	91	92	91	93	91	90	*91.7*	*90.4*
d = 20	Test '$Y \to X$' %	91	90	89	87	91	90	87	88	90	90	*89*	
PLSrc	Test '$X \to Y$' %	94	94	94	94	94	94	96	96	94	96	*94.4*	*94.5*
d = 30	Test '$Y \to X$' %	94	92	94	96	96	94	93	96	94	94	*94.6*	

As we can see, the average value of mutual recognition rate for $Y \leftrightarrow X$ reaches 93.7% for CCArc and 94.5% for PLSrc, using optimal value of parameter d for these methods. Experiments are conducted for assessing the mutual retrieval of the different categories of images, vs. the number of image pairs (classes) in database of FaReS. In the frame of CCArc, the experiment model is the following:

$$\text{DBVIS/Th}(2 \div K/(1(X + N)+1(Y))/1\{\text{CCArc}:320 \times 240 \to (20 \times 20)/\text{MDC/L}_1/rank = 1\}.$$

Figure 10 presents the results obtained, with ordinate axis showing the correct classification rates (in %) and abscissa axis showing the value of parameter 'K' – number of image pairs VIS/Th (i.e., classes) in database of FaReS. Symbol 'Δ' denotes averaged result for 10 variants of input data generation. As it is seen, results of mutual recognition of VIS and Th images for $K > 8$ fall within the range of 91–96%, and average value close to 93%, and agree with the result of previous experiment.

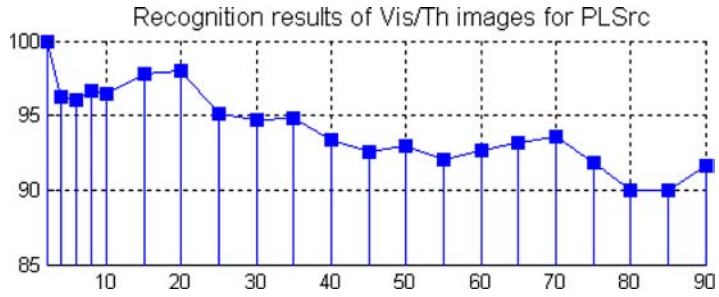
Figure 10 Results of mutual recognition of VIS and Th images for CCArc, vs. the number of classes in database FaReS

For PLSrc method, the experiment model has the following form, corresponding with the above-presented description:

$$\text{DBVIS/Th}(2 \div K/(1(X + N) + 1(Y))/1)\{\text{PLSrc}:320 \times 240 \to (20 \times 20)/\text{MDC/L}_1/rank = 1\}.$$

Figure 11 presents results obtained, where ' ' denotes averaged results of ten experiments. We can see that recognition results are not lower than 95% for $K \leq 35$ and 90% for $35 < K \leq 90$, and average result close to 94%. These results agree with results of previous experiment.

Figure 11 Results of mutual recognition of images VIS and Th vs. the number of classes in database FaReS, for PLSrc (see online version for colours)



6 Conclusions

In the paper, we have presented comprehensive algorithms for realising CCA and PLS and their novel implementations, focusing on the digital image matching. These methods are based on representation of images as sets of rows and columns as their input data. Hence, in analogy to Kukharev and Forczmański (2004), Kukharev and Kuzminski (2003), we have called these methods CCArc and PLSrc. Unlike Ho and Seungjin (2007) and Cai-rong et al. (2007), these methods do not require images to be down-sampled, are direct (non-iterative) contrary to Kukharev and Forczmański (2004), and operate along two directions – rows and columns of input image matrix, making it different than Cai-rong et al. (2007). Size of covariance matrix used in CCArc and PLSrc equals to $DIM = \max\{M, N\}$, allowing practical solution to eigenvalue problem, and ensuring its numerical stability, also for images conforming the requirement of the standard (www.icao.int/mrtd/download/technical.cfm), as concerns the image size. Apart from this, we have shown that the SSS problem is avoided in these methods, because in place of image of size $M \times N$ we actually use N images of size $M \times 1$ and M images of size $N \times 1$. This representation of input data always ensures meeting the condition: $DIM = \max\{M, N\} < (M + N)$.

The algorithms presented in this paper belong to a special group of recognition methods, since they allow to compare images taken from different sources, belonging to different classes (in the aspect of imaging process) or even groups of data. There are no similar algorithms described in the literature. Of course, in case of experiments described earlier, it is still possible to employ any sort of pre-processing to extract features from images and then make a 'normal' comparison, however it changes the whole approach. On the other hand, there are several algorithms and implementations of PLS/CCA, but the methodology of the experiments is not described in a way that is easy to repeat and verify. There is also no test bed for experiments (i.e., databases that integrate multi-modal biometric data), which makes the problem even more complicated. In our opinion, presented accuracy reaching more than 90% makes the method a serious candidate for practical implementation in biometric systems. Moreover, obtained experimental results have confirmed such high effectiveness of presented methods in biometric applications, specifically for processing the face image databases.

The experiments were performed on Equinox database, which included images from visual and thermal bands, which cannot be compared using any traditional recognition algorithms. Moreover, the two-dimensional approach gives better results than

one-dimensional one. On the other hand, the algorithms of PLS/CCA can be employed to solve a rather untypical task of matching images belonging to two different classes (i.e., husband and wife). It was shown on the example of Feret database.

Finally, we have presented a generalised structure of calculations, which can be utilised in any two-dimensional-based projection method.

References

Alın, A., Kurt, S., McIntosh, A.R., Öniz, A. and Özgören, M. (2009) 'Partial least squares analysis in electrical brain activity', *Journal of Data Science*, Vol. 7, pp.99–110.

Cai-rong, Z., Ning, S., Zhen-hai, J. and Li, Z. (2007) '2DCCA: a novel method for small sample size face recognition', *IEEE Workshop on Application of Computer Vision, WACV'07,* Austin, Texas, pp.43–47.

Donner, R., Reiter, M., Langs, G., Peloschek, P. and Bischof, H. (2006) 'Fast active appearance model search using canonical correlation analysis', *IEEE Transaction on PAMI*, Vol. 28, No. 10, October, pp.1960–1964.

Forczmanski, P. and Kukharev, G. (2007) 'Comparative analysis of simple facial features extractors', *Journal Real-Time Image Processing*, Vol. 1, No. 4, pp.239–255.

Ho, L.S. and Seungjin, C. (2007) 'Two-dimensional CCA', *IEEE Signal Processing Letters*, Vol. 14, No. 10, October, pp.735–738.

Hotelling, H. (1936) 'Relations between two sets of variates', *Biometrika*, Vol. 28, pp.321–377.

Jing, X-Y., Wong, H-S. and Zhang, D. (2006) 'Face recognition based on 2D Fisherface approach', *Pattern Recognition*, Vol. 39, pp.707–710.

Kukharev, G. and Forczmanski, P. (2001) 'Hierarchical method of reduction of features dimensionality for image recognition and graphical data retrieval', *Sixth International Conference on 'Pattern Recognition and Information Processing' (PRIP'01)*, 15–17 May, Minsk, Belarus, Vol. 2, pp.57–72.

Kukharev, G. and Forczmański, P. (2004) 'Data dimensionality reduction for face recognition', *Journal Machine Graphics & Vision*, ISSN 1230-0535, Vol. 13, Nos. 1–2, pp.99–121.

Kukharev, G. and Forczmanski, P. (2005) 'Face recognition by means of two-dimensional direct linear discriminant analysis', *PRIP'05*, Minsk, Republic of Belarus, Vol. 2, pp.63–67.

Kukharev, G. and Forczmanski, P. (2007) 'Facial images dimensionality reduction and recognition by means of 2DKLT', *Journal Machine Graphics & Vision*, Vol. 16, Nos. 3–4, pp.401–425.

Kukharev, G. and Kamenskaya, E. (2009) 'Two-dimensional canonical correlation analysis for face image processing and recognition', *Metody Informatyki Stosowanej*, Vol. 18, No. 3, pp.103–112.

Kukharev, G. and Kuzminski, A. (2003) *Biometric Technology: Face Recognition Methods*, Pracownia Poligraficzna WI PS, Szczecin (Poland) (in Polish).

Nhat, V.D.M. and Lee, S.Y. (2007) 'Image-based subspace analysis for face recognition', in Delac, K. and Grgic, M. (Eds.): *Face Recognition*, I-Tech, Vienna, Austria, pp.321–336.

Philips, P.J., Wechler, H., Huang, J. and Rauss, P. (1998) 'The FERET database and evaluation procedure for face recognition algorithms', *Image and Vision Computing*, Vol. 16, No. 5, pp.295–306.

Shan, C., Gong, S. and McOwan, P.W. (2008) 'Fusing gait and face cues for human gender recognition', *Neurocomputing*, Vol. 71, Nos. 10–12, pp.1931–1938.

Sun, L., Ji, S.W., Yu, S.P. and Ye, J.P. (2009) 'On the equivalence between canonical correlation analysis and orthonormalized partial least squares', *Proceedings of the 21st International Joint Conference on Artificial Intelligence*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, http://ijcai.org/papers09/Papers/IJCAI09-207.pdf

Szaber, M. and Kamenskaya, E. (2008) 'Face recognition systems for visible and infrared images with application of CCA', *(in Polish) Metody Informatyki Stosowanej*, Vol. 16, No. 3, pp.223–236.

Tsapatsoulis, N., Alexopoulos, V. and Kollias, S. (1998) 'A vector based approximation of KLT and its application to face recognition', *Proc. IX European Signal Processing Conf. EUSIPCO'98*, Greece, Vol. III, pp.1581–1584.

Wegelin, J.A. (2000) *A Survey of Partial Least Squares (PLS) Methods, with Emphasis on the Two-Block Case*, Technical Report No. 371, University of Washington, March 19.

Wold, S. (1966) 'Nonlinear estimation by iterative least squares procedures', in David, F.N. (Ed.): *Research Papers in Statistics*, pp.411–444.

Yang, J., Zhang, D., Frangi, A.F. and Yang, J-Y. (2004) 'Two-dimensional PCA: a new approach to appearance-based face representation and recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 1, pp.131–137.

Yang, M-L., Sun, Q-S. and Xia, D-S. (2008) 'Two-dimensional partial least squares and its application in image recognition', in Huang, D-S., Wunsch, D.C., Levine, D.S. and Jo, K-H. (Eds.): *ICIC 2008, CCIS 15*, Springer-Verlag, Berlin Heidelberg, pp.208–215.

Ye, J. (2005) 'Generalized. low rank approximations of matrices', *Machine Learning*, Vol. 61, Nos. 1–3, pp.167–191.

Ye, J., Janardan, R. and Li, Q. (2005) 'Two-dimensional linear discriminant analysis', in Saul, L.K., Weiss, Y. and Bottou, L. (Eds.): *Advances in Neural Information Processing Systems 17*, MIT Press, Cambridge, MA, pp.1569–1576.

Yi, D., Liu, R., Chu, R.F., Lei, Z. and Li, S.Z. (2007) 'Face matching between near infrared and visible light images', *LNCS*, Vol. 4642, pp.523–530.

Websites

ISO/IEC JTC 1/SC 37 N 506, *Biometric Data Interchange Formats*, Part 5: Face Image Data – www.icao.int/mrtd/download/technical.cfm

The Equinox IR Face database, http://www.equinoxsensors.com/products/HID.html