Control and Cybernetics

vol. **40** (2011) No. 3

Application of cascading two-dimensional canonical correlation analysis to image matching*†

by

Paweł Forczmański[1], Georgy Kukharev[1,2] and Ekaterina Kamenskaya[2]

[1]West Pomeranian University of Technology,
Faculty of Computer Science and Information Technology
Żołnierska 49, 71–210 Szczecin, Poland

[2]St.-Petersburg Electrotechnical University,
Department of Computer Software Environment
Professora Popova str., 197376 St. Petersburg, Russia

e-mail: {pforczmanski, gkukharev}@wi.zut.edu.pl, ekamenskaya@google.com

Abstract: The paper presents a novel approach to Canonical Correlation Analysis (CCA) applied to visible and thermal infrared spectrum facial images. In the typical CCA framework biometrical information is transformed from original feature space into the space of canonical variates, and further processing takes place in this space. Extracted features are maximally correlated in canonical variates space, making it possible to expose, investigate and model latent relationships between measured variables. In the paper the CCA is implemented along two directions (along rows and columns of pixel matrix of dimension $M \times N$) using a cascade scheme. The first stage of transformation proceeds along rows of data matrices. Its results are reorganized by transposition. These reorganized matrices are inputs to the second processing stage, namely basic CCA procedure performed along the rows of reorganized matrices, resulting in fact in proceeding along the columns of input data matrix. The so called cascading 2DCCA method also solves the Small Sample Size problem, because instead of the images of size $M \times N$ pixels in fact we are using N images of size $M \times 1$ pixels and M images of size $1 \times N$ pixels. In the paper several numerical experiments performed on FERET and Equinox databases are presented.

Keywords: canonical correlation analysis, image matching, face recognition.

*Submitted: September 2010; Accepted: August 2011.

†This is an extended and amended version of the paper, presented at the 5th Congress of Young IT Scientists (Międzyzdroje, 23-25.IX.2010).

1. Introduction

The task of automatic human authentication and identification on the basis of facial portraits is performed by a face recognition system, which in most cases contains initial processing, face detection, feature extraction, and classification or comparison stages. In the scientific literature face recognition systems are classified into three categories (Zhao et al., 2000) based on the way they represent and use face information, namely as: 3D models of frontal part of head (Scheenstra et al., 2005), anthropometric face parameters like graph models (Pentland, 2000) and elastic 2D face models (Lanitis, 1997), and finally, certain set of primitive features (physical or mathematical), mainly distances and angles (Kompanets et al., 2002). In practical realizations, in systems belonging to the third category the original (low-level) features of a face image are usually represented by the luminance value of every pixel. This is motivated by the following observations:

- brightness features are a natural representation of digital images, even after scaling or rotating (Zhao et al., 2000);
- it is possible to find areas with strong changes of brightness (gradients, leaps) and assign them to specific face areas like eyes or pupils, corners of eyes, eyebrows, nose, lips, hair-line, bottom part of face oval; estimation of these points is one of the first steps in 3D model construction (Tistarelli and Grosso, 1997);
- by using brightness of pixels it is possible to build facial models based on facial image approximation in the *eigenface* space (Turk and Pentland, 1991; Kukharev and Kuzminski, 2003).

Recently, experts in biometric technologies paid more attention to Canonical Correlation Analysis (CCA) (Donner et al., 2006), a method providing alignment of different observations of the same set of people from seemingly unrelated measurement sources or points of view. Thus, the CCA can be used in the case of two views of the same semantic object in order to extract the representation of the semantics, i.e. in speaker recognition, when we consider audio signal and lip movements or in image retrieval involving pure image features (color, texture) and associated text (set of tags). In CCA applied to human identification, biometrical information concerning each person (and group of people) is transformed from original space of features into the space of canonical variates, then all further processing takes place in this space (Cai-rong et al., 2007). The most important is that the transformed features are maximally correlated in canonical variates space, making it possible to expose, investigate and model latent relationships between measured variables (Borga, 2001).

In this paper we focus on the case when original features are taken as pixel intensities acquired during two independent observation procedures or taken from two different imaging sources. Further we denote such features as X and Y, respectively. As an example, Fig. 1 illustrates three possible variants of generating

X and Y sets, where FE are corresponding feature extraction units, gathering pixel intensities in the visible light ($FE1$ in variants 1, 2 and 3), positions of markers in the thermal spectrum ($FE2$ in variant 2), and dynamics of motion ($FE2$ in variant 3).

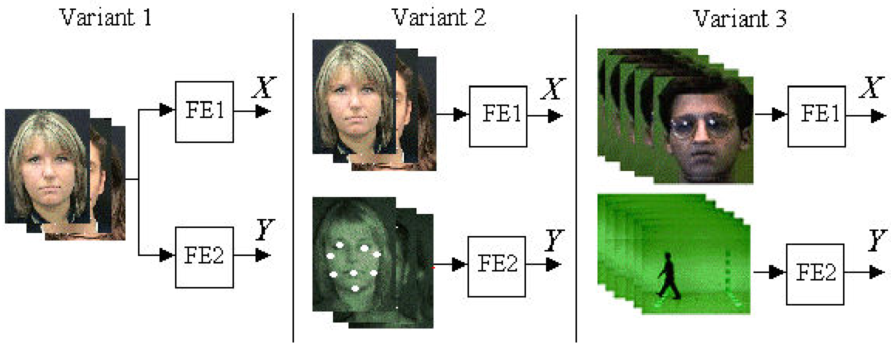

Figure 1. Variants of input data collecting

As mentioned above, an original feature space is constituted by pixel intensities in the image matrix. The dimensionality of such feature space is often very high (related to image size), hence it is very difficult to employ directly CCA to these data. In this paper a new solution to this problem is presented.

We define a task of retrieval of an image from set Y based on given image from set X or in reverse, an image from set X based on given image from set Y, in the sense that each pair $\{X^{(k)}, Y^{(k)}\}$ defines a single object in the database (i.e. one semantic object or a person). The pair $\{X^{(q)}, Y^{(q)}\}$ is a query. On the basis of $X^{(q)}$ (or $Y^{(q)}$) we want to retrieve respective image $X^{(k)}$ (or $Y^{(k)}$) from the database. Resulting image is then used to find an image of the same class, yet belonging to different category. In this way we solve the task of mutual recognition $X \leftrightarrow Y$. The CCA method is especially suitable for this kind of tasks and for images belonging to two independent datasets and two different categories. It should be also noticed that in the case of images we need to use two-dimensional CCA.

The most important parameters in all image processing and pattern recognition operations are the size of image, namely M and N pixels, and the number of images K in the database. In general, the dimension of original feature space $DIM = M \times N$ is greater than the number of images ($DIM >> K$), leading to the problem known as Small Sample Size Problem (SSS) (see Cai-rong et al., 2007). Hence, in order to make application of CCA to image processing tasks efficient, it is required to meet the condition:

$$DIM < K. \tag{1}$$

There are two main approaches to the CCA applied to image processing in biometry that ensure the fulfillment of condition (1) and enable to apply the so called basic CCA method (see Borga, 2001) which was originally developed for vectorized (1D) data only, meaning the necessity to transform inherently two-dimensional matrix (2D) image data to 1D form and to arrange all data points in columns of a data matrix. First approach (Donner et al., 2006, and Yi et al., 2007) is based on transformation of original image into one-dimensional feature vector, with size $DIM_{new} < M \times N$, and application of basic CCA procedure to resulting one-dimensional data. Second approach (Cai-rong et al., 2007; Lee et al., 2007; Kukharev and Kamenskaya, 2009) is based on treating original image as a collection of rows and columns (1D data) and application of basic CCA procedure to these collections. In order to meet condition (1), Lee and Choi (2007) first used down-sampling procedure to reduce the resolution of original data, and then performed CCA. Cai-rong et al. (2007) also used down-sampling for original images and reduced the resolution of images treating them as a collection of columns (1D data), for which basic CCA procedure can be used. This approach is called one-directional CCA. In Cai-rong et al. (2007), Lee and Choi (2007) the procedure of down-sampling was used explicitly or implicitly to meet condition (1). Kukharev and Kamenskaya (2009) presented direct two-dimensional CCA method without preliminary down-sampling. It was based on independent application of CCA to the collection of rows and columns of the set of K images, i.e. solving two individual eigenproblems, yielding two orthogonal eigenvector sets (one for collection of rows and one for collection of columns) organized in two transformation matrices. By multiplying the set of original images from the left and right sides by these matrices we get a representation in canonical variates space. Besides, the authors solved the problem of SSS, since instead of individual images of size $M \times N$ we have N images of size $M \times 1$ and M images of size $N \times 1$. In such form, each image is represented by a set of N lines of $M \times 1$ pixels, and M lines of $1 \times N$ pixels, respectively. The graphical representation of this operation is depicted in Fig. 2.

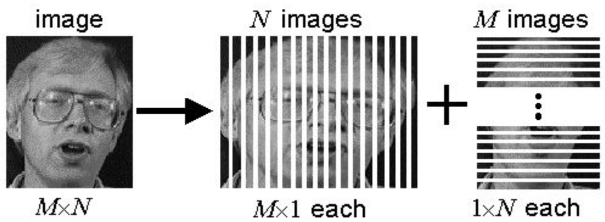

Figure 2. Changing image from matrix into vector form

Such representations of original data always ensure that:

$$DIM_{new} = max\{M, N\} << (M + N)K. \quad (2)$$

The method presented by Kukharev and Kamenskaya (2009) is called CCArc, stressing the fact that the basic CCA method is applied two times - along rows (r) and, separately, along columns (c).

The aim of this paper is to present a novel method of CCA calculation, also based on representation of original images as collection of rows or as collection of columns. Unlike CCArc presented in Kukharev and Kamenskaya (2009), here we use a two-step CCA procedure arranged in the cascade scheme. First, CCA is applied to the collection of rows, then it is applied to the resulting set along its columns (or in reverse order, first to columns, and then to rows). In each step we obtain a matrix of projections. Thus, in result the original images are transformed by the two-directional transformation into two-dimensional canonical variate *images*. The novel method retains all merits of CCArc. Furthermore, new bases are better adapted to the input data characteristics. This ensures wide applicability of the method in various image processing tasks, including biometric applications where condition (1) is mostly violated. In the paper this new algorithm is called *cascading 2DCCA*, since it is oriented at cascading calculations and processing of images as two-dimensional mathematical structures. Hence, the paper is structured as follows. Section 2.1 is devoted to presentation of one-dimensional CCA computation. Section 2.2 presents two-dimensional CCA (CCArc) and the structure of cascading 2DCCA computation. Section 2.3 describes feature space dimensionality reduction achieved by 2DCCA procedure. Sections 3.1 and 3.2 present results of experiments with subsets of images from two well known face image databases: FERET and Equinox. Section 4 enumerates the characteristic features of cascading 2DCCA. Section 5 offers conclusions drawn from the theoretical analysis and conducted experiments.

2. Algorithm outline

2.1. One-dimensional CCA calculation structure

We are given two sets of data, X and Y, each consisting of K vectors of size $DIM \times 1$ each:

$$\begin{cases} X = \left\{X^{(1)}, X^{(2)}, \ldots, X^{(K)}\right\} \\ Y = \left\{Y^{(1)}, Y^{(2)}, \ldots, Y^{(K)}\right\} \end{cases}, \tag{3}$$

so that both X and Y have sizes equal to $DIM \times K$.

The aim of original CCA is to find two projection matrices W_x and W_y transforming input data into the canonical variates spaces U and V, satisfying the criterion (see Hotelling, 1936):

$$\|U - V\| \rightarrow min, \tag{4}$$

where $U = {W_x}^T X$ and $V = {W_y}^T Y$. Corresponding projection matrices are the solutions of the two following eigenproblems (see Hoteling, 1936, and Borga, 2001):

$$\begin{cases} \left(C_{xx}^{-1}C_{xy}C_{yy}^{-1}C_{yx}\right)W_x = W_x\Lambda_x \\ \left(C_{yy}^{-1}C_{yx}C_{xx}^{-1}C_{xy}\right)W_y = W_y\Lambda_y \end{cases}, \tag{5}$$

where: $C_{xx}, C_{xy}, C_{yy}, C_{yx}$ - covariance matrices of input data; Λ_x, Λ_y – diagonal matrices of eigenvalues; W_x, W_y – projection matrices (matrices of eigenvectors); T and -1 denote the operations of matrix transposition and inversion, respectively (all matrices have size DIM).

Computational complexity of (5) is related to DIM. The solution of the eigenproblems depends on values of the parameters DIM and K as well as the condition (1). Meeting that condition is particularly important in applications involving digital images, because DIM is determined by the number of rows and columns of images (i.e. number of pixels), which is usually much greater than the number of images K.

2.2. Two-dimensional CCA

Let us consider two sets X and Y, consisting of K images each, with zero mean and size $M \times N$ pixels, so that:

$$\begin{cases} X = \left\{X^{(1)}, X^{(2)}, \ldots, X^{(K)}\right\} \\ Y = \left\{Y^{(1)}, Y^{(2)}, \ldots, Y^{(K)}\right\} \end{cases}, \tag{6}$$

where $M \times N >> K$ and both X and Y have sizes equal to $M \times (KN)$. Each k-th semantically synonymous object is defined by a pair $\{X^{(k)}, Y^{(k)}\}$. The correlation between variates X and Y is very low in their original coordinate system (original features space). In 2DCCA we are looking for four projection matrices W_{x1}, W_{y1}, W_{x2} and W_{y2} in order to make two-directional projection of each pair $\{X^{(k)}, Y^{(k)}\}$ onto the spaces of canonical variates, so that $X^{(k)} \to U_2^{(k)}$, $Y^{(k)} \to V_2^{(k)}$ and $\|U_2 - V_2\| \to min$ for $k = 1, 2, \ldots, K$. The lower index of U_2 and V_2 was introduced to emphasize the difference with the one-dimensional case of CCA calculation. Thus, after transformation into a new coordinate system, the correlation between U_2 and V_2 is maximized.

In order to assess the characteristic features of such new 2DCCA method, we compare it with CCArc, by presenting its flow of computations (see Kukharev and Kamenskaya, 2009):

$$\begin{cases} \{X, Y\} \xrightarrow{\text{CCAr}} \{W_{x1}, W_{y1}\}, \\ \{X, Y\} \xrightarrow{\text{R}} \{X_1, Y_1\}, \\ \{X_1, Y_1\} \xrightarrow{\text{CCAr}} \{W_{x2}, W_{y2}\}, \\ \left\{ \begin{array}{l} U_2^{(k)} = W_{x1}^T X^{(k)} W_{x2} \\ V_2^{(k)} = W_{y1}^T Y^{(k)} W_{y2} \end{array} \right\}. \end{cases} \qquad \left\| U_2^{(k)} - V_2^{(k)} \right\| \to min \tag{7}$$

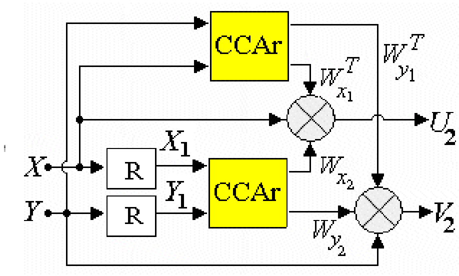

Figure 3. Flow diagram of CCAr

The flow of computations can be observed in Fig. 3. In the first stage of (7) a basic CCA procedure in the direction of image rows is applied (i.e. we treat each image as a collection of its rows) – and consequently we designate this step as CCAr. In the second stage of (7) we reorganize the data (R) by transposition of all image matrices $X^{(k)}, Y^{(k)}$ leading to new sets, denoted X_1 and Y_1. Third stage of (7) implements CCA procedure along rows of X_1 and Y_1 (CCAr), which are in fact columns of X and Y. Results of the first and the third stage are four projection matrices W_{x1} and W_{y1}, W_{x2} and W_{y2}. Final stage of (7) is a transformation of all input images into the canonical variates spaces U_2 and V_2:

$$\begin{cases} U_2 = \left\{U_2^{(1)}, U_2^{(2)}, \ldots, U_2^{(K)}\right\} \\ V_2 = \left\{V_2^{(1)}, V_2^{(2)}, \ldots, V_2^{(K)}\right\} \end{cases} . \tag{8}$$

On the other hand, the new 2DCCA method executes above calculations in the so called cascade scheme presented in Fig. 4.

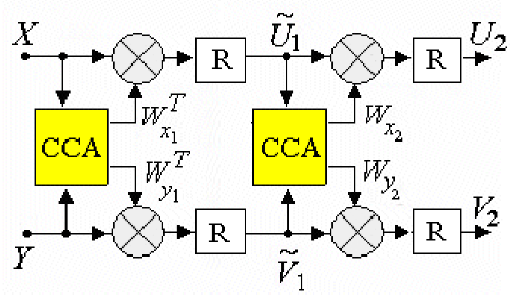

Figure 4. Flow diagram of cascading 2DCCA

The first step involves basic CCA computations along the rows of input matrices X and Y. Projection matrices W_{x1} and W_{y1} are obtained from this step, used to project input data onto the spaces of canonical variates, thus forming the sets U_1 and V_1. In the sets U_1 and V_1 each block of variates of size $M \times N$ is a result of projection of each input image. Second step starts with reorganization of sets U_1 and V_1, where unit R performs transposition of all block of size $M \times N$ contained in U_1 and V_1, resulting in creation of new data sets $\tilde{U}_1$ and $\tilde{V}_1$. Next, the basic CCA procedure is carried out along the rows of reorganized data sets: $\tilde{U}_1$ and $\tilde{V}_1$. Result is a pair of new projection matrices W_{x2} and W_{y2}, used to project sets $\tilde{U}_1$ and $\tilde{V}_1$ onto new canonical variates spaces, denoted $\tilde{U}_2$ and $\tilde{V}_2$, respectively. Because data in $\tilde{U}_1$ and $\tilde{V}_1$ are transposed, we actually carry out CCA procedure along columns of matrices resulting from the first step. Final step is a reorganization of canonical variates contained in sets $\tilde{U}_2$ and $\tilde{V}_2$, resulting in sets U_2 and V_2, where each matrix corresponds to respective input matrix, i.e. X and Y. As a result we get four projection matrices: W_{x1} and W_{y1}, W_{x2} and W_{y2} and two sets of canonical variates (8) with structure corresponding to the structure of input data. In a compact form the model of cascading 2DCCA can be presented as the following sequence of computations:

$$\{X, Y\} \xrightarrow{\text{CCAr}} \{U_1, V_1\} \xrightarrow{\text{R}} \{\tilde{U}_1, \tilde{V}_1\} \xrightarrow{\text{CCAr}} \{\tilde{U}_2, \tilde{V}_2\} \xrightarrow{\text{R}} \{U_2, V_2\}. \tag{9}$$

Individual matrices as results of 2DCCA, corresponding to individual input matrices (images), can be represented as follows:

$$\begin{cases} U_2^{(k)} = W_{x1}^T X^{(k)} W_{x2} \\ V_2^{(k)} = W_{y1}^T Y^{(k)} W_{y2} \end{cases} . \tag{10}$$

Equation (10) clearly exposes the two dimensional nature of projecting the input data, thus distinguishing it from the classical approach aimed at vectorized data. Because these projections are performed as 2D transformations in the basis of eigenvectors, they can be seen as two-dimensional Karhunen-Loeve transformations (2D KLT).The whole procedure is written down as follows:

$$\{X, Y\} \to \frac{\text{2DCCA/2DKLT}}{W_{x1}, W_{x2}, W_{y1}, W_{y2}} \to \{U_2, V_2\}\,. \tag{11}$$

2.3. Feature space dimensionality reduction

In order to achieve dimensionality reduction in (10) and make efficient recognition possible, we reduce the number of eigenvectors using only those which correspond to d greatest eigenvalues (see Kukharev and Kamenskaya, 2009). From $[W_{x1}]^T$ and $[W_{y1}]^T$ we select d rows, corresponding to d greatest eigenvalues, and create dimensionality reduction matrices F_{x1} and F_{y1} . In the matrices W_{x2} and W_{y2} we select d columns, corresponding to d greatest eigenvalues, and create matrices F_{x2} and F_{y2}. Now, *truncated* 2D KLT operation can be written as follows:

$$\begin{cases} \hat{U}_2^{(k)} = F_{x1}X^{(k)}F_{x2} \\ \hat{V}_2^{(k)} = F_{y1}Y^{(k)}F_{y2} \end{cases} . \tag{12}$$

In (12), F_{x1}, F_{y1} and F_{x2}, F_{y2} are matrices of size $d \times M$ and $N \times d$, respectively (or $d_1 \times M$ and $N \times d_2$ in general); symbol ˆ designates difference from (10). Note that $d < min\{M, N\}$ or $d_1 < M$; $d_2 < N$, if $d_1 \neq d_2$. Lower bound of parameter d is selected experimentally, while the upper bound is set according to the typical scheme of selecting eigenvalues in the eigenvector decomposition task. In most cases it employs a criterion, which assumes that the energy of selected components is much higher that the energy of the remainder. Resulting matrices $\hat{U}^{(k)}, \hat{V}^{(k)}$ in (12) generally have dimensions $d_1 \times d_2$ and represent images in canonical variates spaces.

2.4. Classification framework

Classification performed in controlled conditions requires an input image of known category ($X^{(q)}$ or $Y^{(q)}$). This image is a query and the system responds with a class number, which points at the respective image belonging to the other category. The retrieval is performed in the whole set of base images and the results are ordered in the decreasing similarity order. A scheme of such controlled classification is presented in Fig. 5. In the scheme, blocks $2DCCA$ are responsible for cascading two-dimensional Canonical Correlation Analysis (according to the algorithm presented here), blocks C represent comparators (the characteristic features of comparator are described in Section 3). Sample results of this kind of retrieval are presented in Section 3.

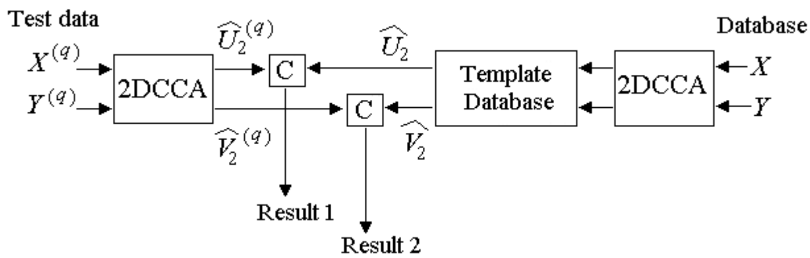

Figure 5. Image retrieval scheme for controlled conditions

However, in practice, we may not know the exact category of the query image, hence we present another scheme of retrieval (see Fig. 6). This scheme is usable in the case of classification of image belonging to the unknown category. The main idea is a fusion (or averaging) of features after 2DCCA of both query and base images. Such fusion can be performed by simple concatenation of feature vectors of images, while the averaging is arithmetic mean of their values. The adequate scheme is presented in Fig. 6.

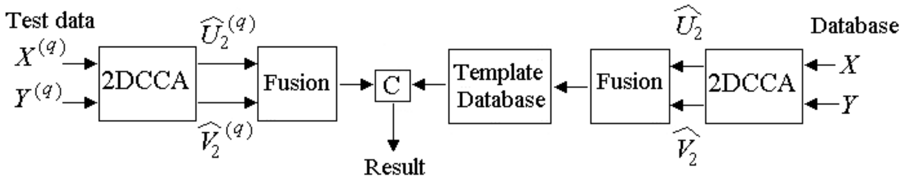

Figure 6. Image retrieval scheme for uncontrolled conditions

3. Experiments

In order to verify the discriminative features of the cascading 2DCCA we performed several experiments on FERET (see Philips et al., 1998) and Equinox (see Equinox, 2009) databases.

3.1. Experiments on images from FERET database

First experiment was conducted for data extracted from FERET database, which contains face images, as presented by Kukharev and Kamenskaya (2009). All images taken from *fa* and *fb* groups have been cropped to face-only area and stored in matrices of 224×184 pixels. Our database consists of two groups of images forming so called "Family album" (see Kukharev and Kamenskaya, 2009). Both groups have the structure as in (5), where $K = 100$. Each pair of images $X^{(k)}$ and $Y^{(k)}$ from one group corresponds to one pair of images from the second group, with different face expression and slight differences in head size and orientation. One of groups is a training set, while images from second group are used for testing (and vice versa). Sample pair of images taken from FERET database is presented in Fig. 7.

Formally, we are solving the task of retrieval of image from set Y based on given image from set X (and vice versa: images from X based on set Y). This retrieval task is formulated as follows: a database contains 100 pairs of test images: 100 images from X and 100 images from Y are used as test images. Retrieval result is regarded as success, if in the first position ($rank = 1$) in the group of retrieval result we have image from the same class as the class of test image. The results of experiment are evaluated as the ratio of correctly classified test images to the total number of test images. As it can be seen from Fig. 8, the standard tool for similarity evaluation, namely phase correlation coefficient, is not suitable for solving this kind of a problem. The picture shows that there is no relation between the two images.

Figure 7. Sample images from fa and fb groups of FERET database (see Philips et al., 2000)

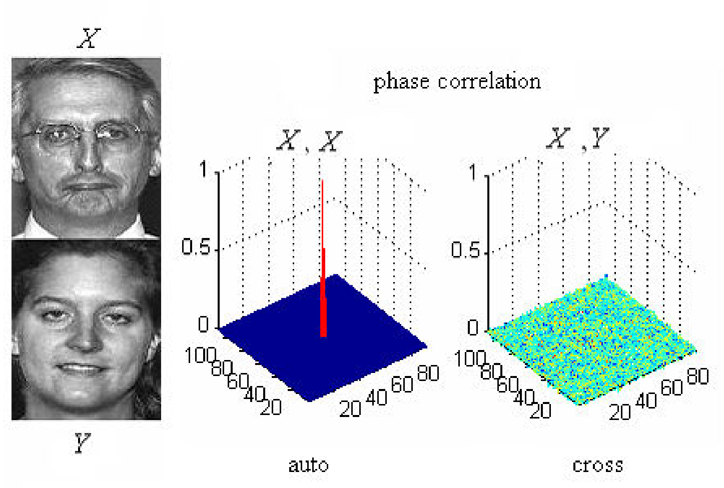

Figure 8. Phase correlation coefficients (autocorrelation and cross-correlation) calculated in the original spatial domain for a pair of images X and Y

In our case we employ the minimal distance classifier (MDC) using the metric L_1 (MDC/L_1). Size of images in canonical variates space is $d \times d$. Hence, the experiment is denoted in accordance to the scheme used by Kukharev and Kamenskaya (2009) as:

$$Feret(100(1_X, 1_Y/1_X, 1_Y))\{224{\times}184 \rightarrow 2DCCA : d{\times}d/MDC/L_1/rank = 1\},$$

where we have 100 classes in the database (templates), each class consisting of one image from X and one image from Y $(1_X, 1_Y)$. Besides, we have 100 test objects, consisting of pairs: one image from set X and one image from set Y $(1_X, 1_Y)$. The images have resolution of 224×184 pixels, given without any pre-processing. The resulting feature space has a dimensionality equal to $d \times d$. Results of retrieval of image X based on image Y (and vice versa) for both methods and values of $d = 5, 6, \ldots, 16$ are shown in Table 1.

Table 1. Recognition accuracy [%] for FERET database

Method	Direction	Recognition accuracy for different d values												
		5	6	7	8	9	10	11	12	13	14	15	16	**mean**
2DCCA	$X \rightarrow Y$	87	89	90	91	91	91	91	91	90	90	89	88	89.9
	$Y \rightarrow X$	85	87	87	88	89	91	91	90	90	88	87	86	88.2
Cascading 2DCCA	$X \rightarrow Y$	89	90	91	92	93	93	93	93	92	92	92	92	91.8
	$Y \rightarrow X$	88	88	90	92	93	93	92	92	92	92	92	92	91.3

As it can be seen, for $d = 10$ we have the best results of recognition for both methods. Data reduction ratio is more than 400 times. The method described in this paper gives better recognition rate by 2-3% in comparison to the former methods. Such increase is significant in case of recognition tasks, when images of male and female faces are "united" in one class. The confirmation of the retrieval ability of the proposed method is presented in Fig. 9. The phase correlation calculated for images in the canonical space of variates gives much higher values for "associated" images than for images from different pairs.

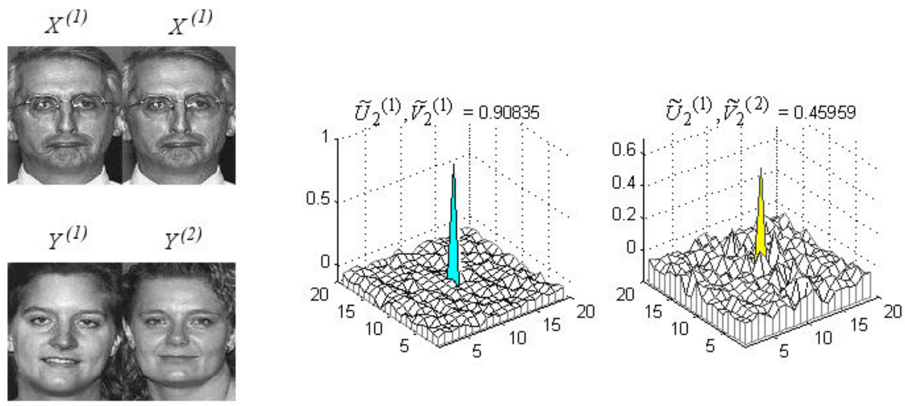

Figure 9. Phase correlation coefficients (for a pair $\{X^{(1)}, Y^{(1)}\}$ and $\{X^{(1)}, Y^{(2)}\}$) calculated in the canonical variates domain

3.2. Experiments on images from Equinox database

In this case the images $X^{(k)}$ and $Y^{(k)}$ for $\forall k \in K$ are taken in visible light (VIS) and thermal band (Th), respectively, from the Equinox database (see Socolinsky

Figure 10. Sample images from Equinox database (Socolinsky and Selinger, 2002)

and Selinger, 2002; Kamenskaya et al., 2009). Here $K = 90$, each image is of size 320×240 pixels, and is kept constant during all experiments. Please note that the size of 320×240 pixels conforms to the smallest allowable size of image face in biometrical applications (see document ISO/IEC, 2009). Each pair of images (*VIS* and *Th*) from the first group corresponds to a pair of images from the second group, having different intensity of background, face expression, small variations in head size and orientation (see Fig. 10). Face images in first group are used for training the algorithm, and images from second group are used for testing (and vice versa). The aim of experiment is retrieval of the image *VIS* based on image *Th*, and vice versa. Formally it means that we will seek the image from set the Y based on given image from set X (and vice versa). The model of experiments can be denoted according to the convention presented above as $Equinox(90(1_X, 1_Y/1_X, 1_Y))\{320 \times 240 \rightarrow 2DCCA : d \times d/MDC/L_1/rank = 1\}$, with the same meaning as described for FERET database. The table below presents obtained classification results.

Table 2. Recognition accuracy [%] for Equinox database

Method	Direction	Recognition accuracy for different d values												
		8	9	10	11	12	13	14	15	16	17	18	19	**mean**
Cascading	$X \rightarrow Y$	98	98	98	98	98	98	98	98	98	98	98	98	98
2DCCA	$Y \rightarrow X$	93	94	96	96	96	96	96	96	96	96	96	96	95.6

It can be seen that recognition rate for $X \rightarrow Y$ test is 98%, and for $Y \rightarrow X$ test is 96% at $d = 10$. Dimensionality reduction ratio in this example is $320 \times 240/100 = 768$ times.

4. Features of the cascading 2DCCA

The here presented two-dimensional method of cascading canonical correlation analysis (2DCCA) has the following features:

- CCA is implemented along two directions (along rows and columns), and consecutive stages of processing are connected in a cascade scheme. Each stage uses basic CCA procedure (see Hotelling, 1936, and Borga, 2001).
- First stage of 2DCCA proceeds along rows of data matrices, and its results are reorganized by transposing all blocks - matrices of dimension $M \times N$. These reorganized matrices are inputs to second processing stage.
- At the second stage of 2DCCA the basic CCA procedure is performed along the rows of reorganized matrices, resulting in fact in proceeding along the columns of input data.
- Obtained results are reorganized again, leading to their form equal to the organization of input data.
- Individual output matrices (images in canonical variates spaces) can be obtained as result of two-dimensional Karhunen-Loeve procedure, in conformance with formula (10) or (12).
- Matrices in (12) have d^2 elements, $\frac{MN}{d^2}$ less than the number of elements in original image matrix. It means that the dimensionality reduction ratio is equal to $\frac{MN}{d^2}$.
- The largest size of covariance matrix in (5) is $max\{M, N\}$, determining the possibility to solve the eigenproblem in (5) and its stability, also for images conforming the ISO/IEC standard.
- Finally, cascading 2DCCA method solves also the SSS problem (see Cairong et al., 2007), where $DIM > K$, because instead of the images of size $M \times N$ in fact we are using N images of size $M \times 1$ and M images of size $1 \times N$. Input data treated in this way always satisfy the condition (2).

5. Summary

The paper presents a new structure and processing scheme of cascading two-dimensional CCA specialized at digital image processing. Characteristic features of the method and its implementation were discussed. Example of application for the task of face image recognition has been shown. In experiments we used the well known benchmark face image databases - FERET and Equinox.

The method presented in the article solves the problem of dimensionality reduction and improves the correlation, as well as allows to compare images from different classes or groups of data. Although there are no similar methods described in the literature, it is still possible to solve such tasks by employing algorithms of pre-processing to extract features from images and then perform a straightforward comparison (understood as a comparison between features of the same class). However, it would change the whole approach and would not be consistent with the main idea of the paper. Despite the fact that there are several algorithms and implementations of CCA described in the literature, the methodology of experiments is not shown 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. Hence, the motivation of the work presented in this paper. Results obtained from experiments showed high efficiency of the proposed method, in comparison to the state-of-the-art 2DCCA presented earlier by Zuo et al. (2009). Provided accuracy of more than 91–98% (depending on the application) makes the method a serious candidate for practical implementation in multi-modal biometric systems.

References

Borga, M. (2001) Canonical correlation - a tutorial. see http://www.imt.liu.se/magnus/cca/tutorial/tutorial.pdf.

Cai-rong, Z., Ning, S. and Zhen-shai, J. (2007) 2DCCA: A novel method for small sample size face recognition. In: *IEEE Workshop on Application of Computer Vision, WACV'07*. IEEE Computer Society Press, Los Alamos, CA, 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*, **28** (10), 1960–1964

Equinox Corporation (2009) Human Identification at a Distance: The Equinox database. see http://www.equinoxsensors.com/products/HID.html.

Hotelling, H. (1936) Relations between two sets of variates. *Biometrika* **28** (3-4), 321–377.

ISO/IEC (2009) JTC 1/SC 37 N 506: Biometric data interchange formats, part 5: Face image data. see http://www.icao.int/mrtd/download/technical.cfm

Kamenskaya, E., Borawski, M. and Szaber, M. (2009) Visible and infrared recognition using canonical variables. *Polish Journal of Environmental Studies* **18** (3B), 39–43.

Kompanets, L. et al. (2002) Biometrical method of personal identification based on information about asymmetry and mimics (in Polish). In: *Proceedings of the 6-th Intern. Conference IT Forum Secure 2002*, 2, NASK and Multicopy Press, Warszawa, 31–40.

Kukharev, G. and Kuzminski, A. (2003) *Biometric techniques. Part I. The Methods of Face Recognition* (in Polish). Pracownia Poligraficzna WI PS, Szczecin (Poland).

Kukharev, G. and Kamenskaya, E. (2009) Two-dimensional canonical correlation analysis for face image processing and recognition. *Metody Informatyki Stosowanej* **18** (3), 103–112.

Lanitis, A., Taylor, Ch.J. and Cootes, T.F. (1997) Automatic interpretation and coding of face images using flexible models. *IEEE Trans. PAMI* **17** (7), 73–755.

Lee, S.H. and Choi, S. (2007) Two-dimensional CCA. *IEEE Signal Processing Letters* **14** (10), 735–738.

Pentland, A. and Choudbury, T. (2000) Face recognition for smart environments. *IEEE Computer* **1**, 50–55.

Phillips, P.J., Wechler, H., Huang J. and Rauss, P.J. (1998) The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing* **16** (5), 295–306.

Phillips, P.J. , Moon, H., Rauss, P.J. and Rizvi, S. (2000) The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (10), 1090–1104.

Scheenstra, A., Ruifrok, A., Veltkamp, R. (2005) A Survey of 3D Face Recognition Methods. In: *Audio- and Video-Based Biometric Person Authentication.* **LNCS 3546**, Springer, 325–345.

Socolinsky, D. and Selinger, A. (2002) A comparative analysis of face recognition performance with visible and thermal infrared imagery. In: *Proceedings of the 16th International Conference on Pattern Recognition.* IEEE Computer Society Press, Los Alamos, CA, 217–222

Tistarelli, M. and Grosso, E. (1997) Active face recognition with a hybrid approach. *Pattern Recognition Letters* **18**, 933–946.

Turk M., Pentland, A. (1991) Eigenfaces for Recognition. *Journal of Cognitive Neurosicence* **3**, 1, 71–86

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

Zhao, W., Chellappa, R., Rosenfeld, A. and Phillips, P. (2000) Face recognition: A literature survey. *ACM Computing Surveys* **35** (4), 399–458.

Zuo, W., Wang, K. and Zhang, H. (2009) Subspace Methods for Face Recognition: Singularity, Regularization, and Robustness. *State of the Art in Face Recognition.* InTech, Vienna, Austria.