Optica Applicata, Vol. XXXII, No. 4, 2002

Application of thermal imaging for human face recognition

JOANNA KOBEL

Bio-Optics Group, Institute of Physics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50–370 Wrocław, Poland.

ARTUR SUCHWAŁKO

Institute of Mathematics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50–370 Wrocław, Poland.

HALINA PODBIELSKA

Bio-Optics Group, Institute of Physics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50–370 Wrocław, Poland.

In this paper the possibility of exploiting thermal images of human faces and their computer aided analysis for person recognition is presented. The statistical pattern recognition methods realizing the classification and identification of human faces thermograms are applied. For thermogram recording the spectral range of 8–13 μm was chosen. The applied procedure is a two step scheme: dimensionality reduction of the data based on the PCA (principal component analysis) approach and cluster analysis. Five hierarchical methods of cluster analysis and five different distance metrics were used for classification. The recognition was carried out for six different sizes of face images: 85×85, 50×50, 25×25, 13×13, 7×7 and 3×3 pixels. It was shown that the cluster analysis by the Ward method gave the best results, namely the highest recognition accuracy (RA). The RA depends on the size of an analyzed image, as well. It decreases with size reduction.

1. Introduction

Face recognition belongs to the most fundamental human perceptual skills. It is possible because each face is unique and unrepeatable, even the faces of twins. For that reason many research centres are conducting intensive investigations in this area. There exist many systems for identification and recognition of human faces. However, the majority is based on ordinary pictures recorded by film or digital cameras. In this paper we propose to use thermal images of human face for person recognition.

Thermovision is a method which relies on a remote and contactless estimation of temperature distribution on the object surface. This method is based on the observation and recording of the distribution of infrared radiation emitted by each body whose temperature is higher than absolute zero. For the measurement purposes the far infrared

is used, mostly in the wavelengths from 3 μm to 5 μm or from 8 μm to 13 μm, what is connected with good infrared radiation transmission through atmosphere in these regions. Nowadays, thermovision is exploited in medicine, building industry, heat engineering, electronics, power plants and in energy transmission, environment protection, as well as in military technology.

One of the newest applications of thermovision is face recognition. This approach is based on the demonstrated premise that thermal image of each human's face is characteristic, like fingerprints, iris and retinal patterns, hand geometry *etc.* This application is possible because temperature distribution of each face is individual. The warmest parts of the face are eye areas, forehead, mouth, while the coolest are the nose and ears [1], [2].

Thermal facial characteristics differs from visible facial features. The vascular system in the human face produces a unique facial pattern when heat passes through the facial tissue and is emitted from the skin. Therefore, this method distinguishes the living subjects from dead ones, what is impossible in the case of pure optical methods in visible spectrum. Thermal recognition of human faces is an external illumination, independent and non-intrusive. The identity can be verified without contact, in low light or even in the absence of light. Moreover, thermograms cannot be altered or camouflaged by masks or cosmetic surgery. However, the external masks can change the thermal pattern, temporarily [3].

The infrared camera captures the patterns of heat emitted by body surfaces, which are referred to as thermograms. Thermal images, like other images of human face, can be read, digitized and stored by a computer. Storage capacity requirements are minimal, since the system only needs a portion of the face for recognition, thereby requiring only a few kilobytes of digitized imagery to be stored.

2. Experimental method

For the measurement we used the thermal system AGEMA 900 LW. The sensing range was chosen in the spectral region 8–13 μm (most popular region for performing the thermal examination of living objects). The experimental system consisted of thermal camera connected to the personal computer (see Fig. 1). The captured pictures were observed on the monitor, preprocessed and stored on the hard disk. In this way, the database of thermograms was created.

The computer preprocessing of recorded images was the first step of thermogram analysis. All thermograms were correspondingly calibrated into the same scale by means of the computer program IRwin Research 2.01 from AGEMA Infrared System. The scale Iron was chosen. The scale is the way in which the pseudocolours are assigned to the certain temperatures on the body surface.

The recorded pictures were in form of 135×270 pixels bitmaps. For each picture the square sample of a central face part was chosen as an image of size 85×85 pixels, containing eyes, nose and mouth region (see Fig. 2). Next, the resizing of each image was performed, so thus to obtain the set of following images: 50×50, 25×25, 13×13,

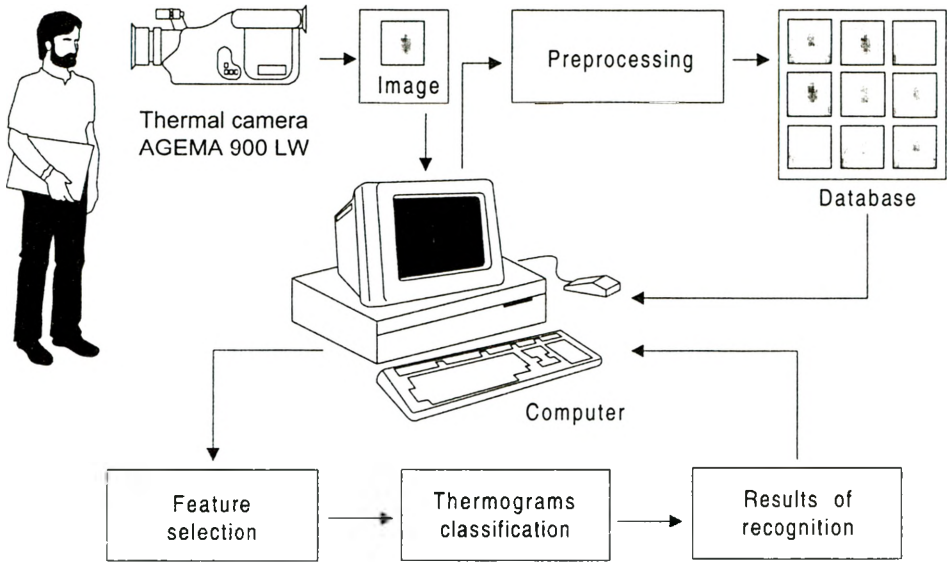

Fig.1 Scheme of the experimental setup used for face thermogram recognition.

7×7 and 3×3 pixels. Then, all images were converted into the feature vectors which were used for further calculations. Each thermogram is a two-dimensional entity, where both dimensions are of spatial nature (x, y). Every point (pixel) on the thermogram has one or more features (*e.g.*, intensity, amplitude), so starting with the two-dimensional array of intensity values $I(x, y)$, we can construct the feature vectors. So, from original $n \times n$ pixels face image, the n^2 element vector $x_i = [x_{i1}, x_{i2}, \ldots, x_{in^2}]^T$ is received. We denoted the spatial resolution of the feature vector by N, where $N = n^2$.

The proper recognition should be a two-step process. The first step is feature selection with the simultaneous dimensionality reduction of data. For that purpose we used the principal component analysis (PCA), also known as Karhunen–Loeve decomposition [4], [5]. Next, the classification based on cluster analysis (CA) was carried out. We used five hierarchical methods: single linkage (SL), complete linkage (CL), average linkage (AL), centroid linkage (CE) and Ward's method (W) (inner

Fig. 2. Original face thermogram (**a**) and central sample of 85×85 pixels size (**b**).

squared distance). Additionally, five methods of distance measure were used: Euclidean (E), standardized Euclidean (SE), cityblock (CB), Mahalanobis (MH) and Minkowski (MI) distance [6]. The results of all methods combination of feature selection, classification and metric types were analyzed. The recognition was realized for six sizes of thermograms.

3. Research material

The experiment was performed on 10 adult healthy volunteers. This group consisted of five females (F) and five males (M), aged 23 to 45. First, thermal images of "natural" human faces (without any covering substances) were recorded. Next, various materials that can influence temperature distribution were put on the faces and new thermal

T a b l e 1. Substances used for covering the faces.

Substance type	Symbol
Ethyl alcohol – 40% concentration	EA
Tonic for normal skin	TN
Tonic for greasy skin	TG
Moisturizing cream	MC
Greasy cream	GC
Loose powder	P
Light liquid make-up	LM
Dark liquid make-up	DM
Male moisturizing cream	MMC
Male vanishing cream	MVC
After shave balm	SB
Cologne	C

T a b l e 2. Recorded images.

	Picture type of substance									
Person	1	2	3	4	5	6	7	8	9	10
1 (F)	Natural	TG	MC	MC15	MMC	MMC15	P	DM	EA	EA15
2 (F)	Natural	TN	MC	MC15	MVC	MVC15	P	LM	EA	EA15
3 (F)	Natural	TG	GC	GC15	MMC	MMC15	P	DM	EA	EA15
4 (F)	Natural	TN	GC	GC15	MVC	MVC15	P	LM	EA	EA15
5 (F)	Natural	TN	MC	MC15	MVC	MVC15	P	LM	EA	EA15
6 (M)	Natural	TG	MC	MC15	MVC	MVC15	SB	SB15	C	C15
7 (M)	Natural	TN	GC	GC15	MVC	MVC15	SB	SB15	C	C15
8 (M)	Natural	TN	MC	MC15	MMC	MMC15	SB	SB15	C	C15
9 (M)	Natural	TG	GC	GC15	MMC	MMC15	SB	SB15	C	C15
10 (M)	Natural	TG	GC	GC15	MMC	MMC15	SB	SB15	C	C15

images were captured. A list of the substances is presented in Tab. 1. Table 2 presents detailed information about recorded thermograms.

All images were recorded immediately after the placement of the covering mask or after 15 minutes (15 – denotes the time in minutes).

The faces were recorded at ambient temperature 21.5 °C and 28% humidity of the environment.

4. Feature selection and classification of thermograms

The feature selection is used for an extraction of the data of these thermograms, which are most suitable for persons' recognition. The attributes that enable proper recognition are generally difficult to define and to be measured directly. Therefore, they have to be received indirectly, with the help of certain transformations of original features space. The new space should have lower dimension and simultaneously hold the sufficient information for realization of a correct classification. For that purpose we choose the PCA, which is perhaps the most popular technique in dimensionality reduction and particularly in the field of face recognition [7]–[10].

4.1. Principal component analysis

Let us to define an ensemble of all M face images as a matrix $D = [x_1, x_2, ..., x_M]$. Size of the matrix D is $N \times M$. Using the mean vector defined as follows:

$$u = \frac{1}{M} \sum_{j=1}^{M} x_i, \tag{1}$$

the modified data set $\hat{D}$, built from zero mean feature vectors (also called caricature vectors)

$$\hat{x}_i = x_i - u \tag{2}$$

is obtained

$$\hat{D} = [\hat{x}_1, \hat{x}_2, ..., \hat{x}_M]. \tag{3}$$

In the next step the covariance matrix of the data $\hat{D}$ set is computed

$$C = \frac{1}{M} \sum_{i=1}^{M} \hat{x}_i \cdot \hat{x}_i^T = \hat{D} \cdot \hat{D}^T \tag{4}$$

as the $N \times N$ matrix.

Eigenvectors of matrix C form the orthonormal basis of the feature subspace. By definition the eigenvector Φ_i and its corresponding eigenvalue λ_i of the covariance matrix are determined by:

$$C \cdot \Phi_i = \lambda_i \Phi_i, \tag{5}$$

$$\Phi_i^T \cdot \Phi_i = 1. \tag{6}$$

For convenience, the eigenvectors and their eigenvalues are then ranked as follows:

$$\lambda_1 > \lambda_2 > \lambda_3 > \ldots \lambda_N. \tag{7}$$

Weighting each vector by the corresponding eigenvalue, we can obtain an estimation of the effectiveness of data description. Moreover, taking a few first eigenvectors, we get an orthonormal reduced space that ensures error and entropy reduction [5], [11], [12].

In many applications of PCA, the covariance matrix becomes too large for practical computations (*e.g.*, when we deal with $M = 100$ images of size 85×85, we have to calculate the eigenvectors of the square matrix of size $N = 7225$). Direct methods of eigenvector calculations are computationally expensive and require $O(N^3)$ operations. In order to overcome this problem, one can apply the following method to calculate first M eigenvectors.

Let us define a new matrix

$$C' = \hat{D}^T \cdot D. \tag{8}$$

Size of the matrix C' is $M \times M$. Let its eigenvectors will be denoted by Φ_i' with their corresponding eigenvalues λ_i'. Assume again that eigenvectors are ordered with respect to their eigenvalues. Eigenvalues of C' are equal to the first M eigenvalues of the original covariance matrix C. Thus we have

$$\lambda_i = \lambda_i'. \tag{9}$$

The original eigenvectors can be calculated by the below formula

$$\hat{\Phi}_i = D \cdot \Phi_i'. \tag{10}$$

After normalization, the unit-norm eigenvectors are as follows:

$$\Phi_i = \hat{\Phi}_i \cdot \frac{1}{\left\| \hat{\Phi}_i \right\|}. \tag{11}$$

Using this efficient implementation, one can get the reduction of matrix size, thus the reduction of the computation time. Even in cases of smaller images, the described method significantly reduces the time of computation.

4.2. Cluster analysis

In the new reduced space, obtained by exploitation of PCA, it is possible to define distances in order to establish a discriminant function that can classify thermograms. Thus, the CA methods for recognition can be used.

Before the clustering, distances between objects d_{ab} are computed (see Tab. 3) and the distances metrics are created.

Next, during the merging of two clusters a and b into a new cluster, the elements of each distance matrix are transformed according to the formula of described below methods [13]. In our experiment five hierarchical methods were applied (see Tab. 4).

T a b l e 3. Distance metrics used in the experiment.

Name of metrics	Mathematical formula
Euclidean	$d_{ab} = \{(x_a - x_b)(x_a - x_b)'\}^{1/2}$
Standarized Euclidean	$d_{ab} = \{(x_a - x_b)D^{-1}(x_a - x_b)'\}^{1/2}$
Mahalanobis	$d_{ab} = \{(x_a - x_b)'V^{-1}(x_a - x_b)\}^{1/2}$
Cityblock	$d_{ab} = \sum_{j=1}^{n} \lvert x_{aj} - x_{bj} \rvert$
Minkowski	$d_{ab} = \left\{ \sum_{j=1}^{n} \lvert x_{aj} - x_{bj} \rvert^{p} \right\}^{1/p}$

T a b l e 4. CA methods used for classification.

Name of method	Mathematical formula
Single linkage	$d_{ab} = \min(d(x_{aj}, x_{bj}))$
Complete linkage	$d_{ab} = \max(d(x_{aj}, x_{bj}))$
Average linkage	$d_{ab} = \frac{1}{n_a n_b} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} d(x_{ai}, x_{bj})$
Centroid linkage	$d_{ab} = d\left(\frac{1}{n_a} \sum_{i=1}^{n_a} x_{ai}, \frac{1}{n_b} \sum_{j=1}^{n_b} x_{bj} \right)$
Ward's method	$d_{ab} = \frac{n_a n_b d^2\left(\frac{1}{n_a} \sum_{i=1}^{n_a} x_{ai}, \frac{1}{n_b} \sum_{j=1}^{n_b} x_{bj} \right)}{n_a + n_b}$

where: d_{ab} – the distance between cluster a and cluster b; n_a, n_b – adequatly, the number of objects in cluster a or b; x_{ai}, x_{bj} – adequatly i-th or j-th object in cluster a or b ($i \in (1, 2, \ldots n_a)$ and $j \in (1, 2, \ldots n_b)$).

In hierarchical clustering, each object is defined like an independent cluster. The algorithm merges two clusters, at every step. Each cluster may be one of the original objects or a cluster merged at some of the previous steps. The choice of which cluster should be merged is based on a measure of the distance between clusters and on a decision rule [6], [14].

5. Results

Our database was composed of thermographic pictures of 10 persons. For each person 10 thermograms were recorded and each thermogram was evaluated taking into account 6 different image sizes. Totally 600 images were analyzed. As a measure of recognition accuracy we used the number of well-classified thermograms (in percentage):

$$\mathrm{RA} = \frac{\sum_{i=1}^{M} N_{\mathrm{wct}}}{\sum_{i=1}^{M} N_i} 100\% \tag{12}$$

where N_{wct} is the number of well-classified thermograms of each class, N_i is the number of samples in the i-th class and M is the number of classes.

Figures 3 and 4 show the analysis results for various CA methods and distance metrics. One can see that the Ward method gave the highest accuracy, independent of the size of the analyzed picture. The worst results were achieved for single linkage.

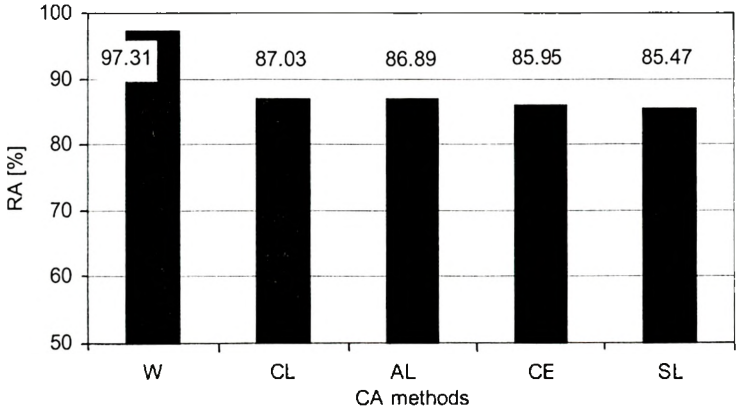

Fig. 3. CA methods ranked from the best to the less accurate.

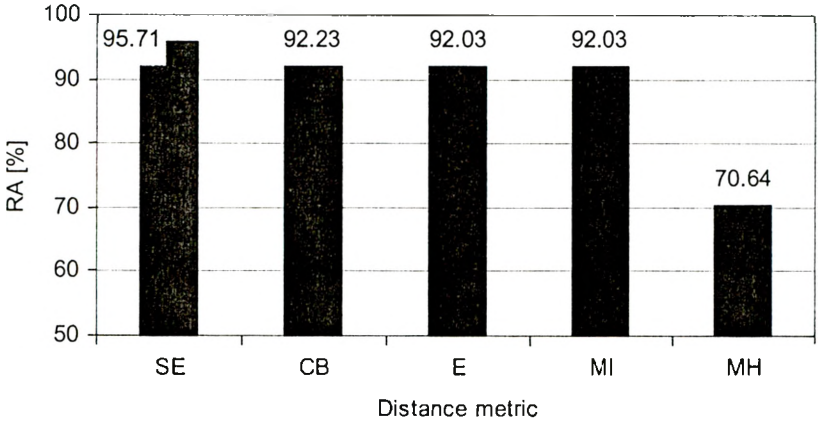

Fig. 4. Distance metrics ranked, starting from the most accurate.

T a b l e 5. Results for the images of 85×85 pixels size.

Methods / Metrice	SL	CL	AL	CE	W	
	Average accuracy [%]					Mean
E	94.79	93.33	95.67	95.78	99.00	95.71
SE	96.78	98.11	98.33	98.11	99.00	98.07
CB	92.89	96.56	95.22	94.89	98.78	95.67
MH	74.11	63.22	65.56	62.22	98.22	72.67
MI	94.78	93.33	95.67	95.78	99.00	95.71
Mean	90.67	88.91	90.09	89.36	98.80	91.57

T a b l e 6. Results for the images of 50×50 pixels size.

Methods / Metrice	SL	CL	AL	CE	W	
	Average accuracy [%]					Mean
E	94.89	92.78	95.44	94.56	99.00	95.33
SE	97.78	97.78	97.00	97.00	99.00	97.71
CB	91.00	96.22	94.11	94.89	98.78	95.00
MH	74.22	62.56	64.00	62.44	98.00	72.24
MI	94.89	92.78	95.44	94.56	99.00	95.33
Mean	90.56	88.42	89.20	88.69	98.76	91.12

As the distance metrics the standardized Euclidean showed the best results. The worst were achieved for Mahalanobis metric. The detailed results are depicted in Tabs. 5–10. There are results for all combinations of CA methods and distance metrics. As one can see, the best result, 99%, was achieved for combination of Ward's method with Euclidean, or standardized Euclidean, or Minkowski metrics for images of size

T a b l e 7. Results for the images of 25×25 pixels size.

Methods / Metrice	SL	CL	AL	CE	W	
			Average accuracy [%]			Mean
E	90.00	92.44	92.89	92.00	98.56	93.18
SE	97.44	97.11	97.00	97.00	98.78	97.47
CB	89.00	95.00	95.11	93.89	98.78	94.36
MH	73.00	62.00	62.33	60.67	97.78	71.16
MI	90.00	92.44	92.89	92.00	98.56	93.18
Mean	87.89	87.80	88.04	87.11	98.49	89.87

T a b l e 8. Results for the images of 13×13 pixels size.

Methods / Metrice	SL	CL	AL	CE	W	
			Average accuracy [%]			Mean
E	85.44	92.11	90.33	89.11	97.67	90.93
SE	91.78	96.11	96.44	96.44	97.56	95.67
CB	88.00	93.56	91.11	91.11	97.00	92.16
MH	68.89	62.11	59.89	60.44	91.78	68.62
MI	85.44	92.11	90.33	89.11	97.67	90.93
Mean	83.91	87.20	85.62	85.24	96.34	87.66

T a b l e 9. Results for the images of 7×7 pixels size.

Methods / Metrice	SL	CL	AL	CE	W	
			Average accuracy [%]			Mean
E	85.33	91.33	92.22	90.67	96.33	91.18
SE	92.89	96.11	95.89	95.67	96.44	95.40
CB	81.78	89.44	92.00	90.56	95.56	89.87
MH	70.00	60.22	65.00	63.11	94.33	70.53
MI	85.33	91.33	92.22	90.67	96.33	91.18
Mean	83.07	85.69	87.47	86.14	95.80	87.63

T a b l e 10. Results for the images of 3×3 pixels size.

Methods / Metrice	SL	CL	AL	CE	W	
			Average accuracy [%]			Mean
E	78.44	87.67	83.89	82.67	96.67	85.87
SE	84.22	92.67	88.33	88.00	96.44	89.93
CB	76.67	90.11	84.44	83.78	96.56	86.31
MH	65.78	62.67	64.11	58.56	92.11	68.65
MI	78.44	87.67	83.89	82.67	96.67	85.87
Mean	76.71	84.16	80.93	79.14	95.69	83.33

85×85. The worst outcome, 58.56%, was for the images of size 3×3 for combination of centroid method and Mahalanobis metric.

6. Discussion

Face recognition based on the thermal imaging still has got little attention compared to the recognition in visible spectrum. Nowadays, there is an urgent need for fast and accurate human identification. The infrared imagery has the advantage of differentiating the living persons from the dead ones, what is not possible when exploiting the optical imaging in visible range.

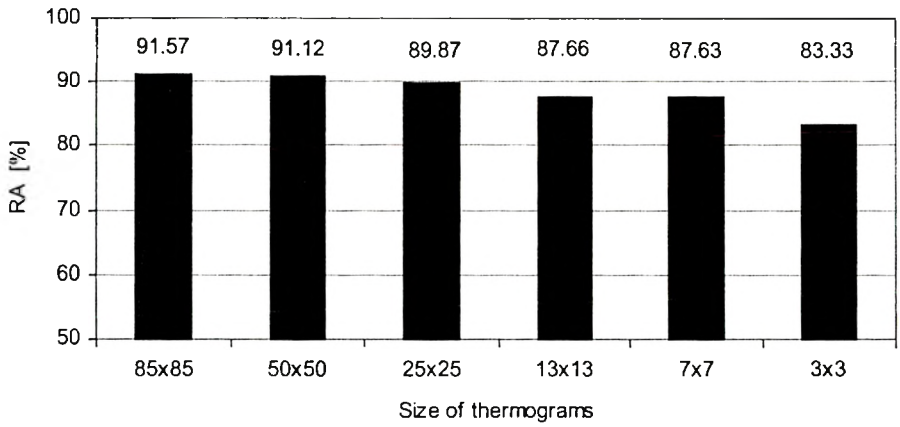

Fig. 5. Mean recognition accuracy for various image sizes.

The analysis can be based even on the reduced-size pictures. Figure 5 show the results depending on the image size. As it was expected, the mean recognition accuracy decreases with the reduction of image size. The best achieved mean accuracy exceeds 91%. However, the worst one was higher than 83%. Practically, such result can be accepted, as well.

Acknowledgments – The authors gratefully acknowledge the support from the State Committee for Scientific Research (KBN), Poland (grant KBN no. 8 T11E 037 19).

References

[1] WILDER J., PHILLIPS P., JIANG C., WIENER S., *Comparison of visible and infrared imagery for face recognition*, [In] Proc. 2nd Inter. Conf. on Automatic Face&Gesture Recognition, Killington 1996, pp. 182–187.

[2] SELINGER A., SOCOLINSKY D., *Appearance-Based Facial Recognition Using Visible and Thermal Imagery: A Comparative Study*, Technical Report 2001–2002, Equinox Corporation, Quebec 2002.

[3] HOŁOWACZ I., KOBEL J., PODBIELSKA H., Proc. SPIE **4356** (2001), 408.
[4] AHMED N., RAO K.R., *Orthogonal Transforms for Digital Signal Processing*, Springer-Verlag, New York 1975.
[5] FUKUNAGA K., *Introduction to Statistical Pattern Recognition*, 2nd edition, Academic Press, San Diego 1990.
[6] EVERITT B., *Cluster Analysis*, 3rd edition, Edward Arnold, London 1993.
[7] SIROVICH L., KIRBY M., J. Opt. Soc. Am. A **4** (1987), 519.
[8] KIRBY M., SIROVICH L., Proc. IEEE Trans. Pattern Analysis and Machine Intelligence, **12** (1990), 103.
[9] TURK M., PENTLAND A., [In] *Proc. IEEE on Computer Vision and Pattern Recognition*, Los Alamos 1991, 586.
[10] TURK M., PENTLAND A., J. Cognitive Neuroscience **3** (1991), 71.
[11] ASH R.B., GARDNER M.F., *Topics in Stochastic Processes*, Academic Press, New York 1972.
[12] FUKUNAGA K., KOONTZ W.L., Proc. IEEE Trans. Comput. **C-19** (1970), 311.
[13] KOBEL J., SUCHWAŁKO A., PODBIELSKA H., ULATOWSKA-JARŻA A., Opt. Eng., in press.
[14] GORDON A.D., *Classification*, Chapman and Hall/CRC, Boca Raton 1980.

Received August 13, 2002