

Tomasz HERUD, Michal KAWULOK  
Silesian University of Technology, Institute of Informatics  
Future Processing, Gliwice, Poland

Bogdan SMOLKA  
Silesian University of Technology, Institute of Automatic Control

## EMOTION RECOGNITION FROM FACIAL IMAGES USING BINARY FACE RELEVANCE MAPS<sup>1</sup>

**Summary.** This paper is focused on automatic emotion recognition from static grayscale images. Here, we propose a new approach to this problem, which combines a few other methods. The facial region is divided into small subregions, which are selected for processing based on a face relevance map. From these regions, local directional pattern histograms are extracted and concatenated into a single feature histogram, which is classified into one of seven defined emotional states using support vector machines. In our case, we distinguish: anger, disgust, fear, happiness, neutrality, sadness and surprise. In our experimental study we demonstrate that the expression recognition accuracy for Japanese Female Facial Expression database is one of the best compared with the results reported in the literature.

**Keywords:** emotion recognition, local binary patterns, face recognition

## ROZPOZNAWANIE EMOCJI NA PODSTAWIE OBRAZÓW TWARZY Z UŻYCIEM BINARNYCH MAP ISTOTNOŚCI

**Streszczenie.** W artykule tym przedstawiono zagadnienie rozpoznawania emocji na podstawie obrazów w skali szarości. Prezentujemy w nim nowe podejście, stanowiące połączenie kilku istniejących metod. Obszar twarzy jest dzielony na mniejsze regiony, które są wybierane do dalszego przetwarzania, z uwzględnieniem binarnych map istotności. Z każdego regionu ekstrahowany jest histogram lokalnych wzorców binarnych, a następnie histogramy są składane do wektora cech i klasyfikowane za pomocą maszyny wektorów podpierających. W naszym przypadku

---

<sup>1</sup> This work was supported by the Polish National Science Center (NCN) under the Grant: DEC-2012/07/B/ST6/01227

rozdzielamy takie emocje, jak: gniew, wstręt, strach, szczęście, neutralność, smutek i zaskoczenie. Podczas naszych eksperymentów pokazaliśmy, że nasze podejście umożliwia poprawę skuteczności rozpoznawania emocji dla bazy Japanese Female Facial Expression względem innych istniejących metod.

**Słowa kluczowe:** rozpoznawanie emocji, lokalne wzorce binarne, rozpoznawanie twarzy

## 1. Introduction

There are six prototypic emotion expressions which are universal across human races and cultures, namely: anger, disgust, fear, happiness, sadness and surprise [1]. Nowadays, emotion recognition from digital still images and from video sequences is more important than ever. Not only is it an object of interest for psychologists, but it may deliver relevant information for the purposes of surveillance, security, for example in airports or train stations. One challenge is the recognition from image sequences, second one is the recognition from static images. The latter is more challenging, because we have less data, therefore less information, but usually it is easier to implement. In case of recognition based on static images, it is important to have an image acquired at an appropriate moment. When someone starts expressing a certain emotion, it is initially hard to determine the emotional state. After a while, the emotion is well manifested and afterwards the expression starts fading away.

In this paper, we present a new approach for recognizing emotions. We calculate local directional pattern (LDP), split image into subimages, then we apply masking using a binary face relevance map (we define the face relevance map as a long vector of binary weights  $m$  which is imposed on a face image when the feature vector is calculated [2]). After that we obtain feature vectors which we use for supervised learning with support vector machines.

The paper is organized as follows. Existing methods are presented in Section 2. Proposed approach is detailed in Section 3. In Section 4, we describe the JAFFE database, which we use for testing our algorithm, and present our results. Section 5 concludes the paper.

## 2. Related literature

In 1990, He and Wang introduced local binary patterns (LBP) – a powerful feature extraction technique for texture classification [3]. LBP is an operator which labels the pixels of an image by assigning a binary number to every neighbor of each pixel and considers the result as a binary number that is assigned to every pixel in the image. LBP was used in many

other works for texture classification [4], face recognition [5, 6] and facial emotion detection [7, 8]. Later, there have been new algorithms introduced based on LBP, like local directional pattern (LDP) [9] or local directional texture pattern (LTDTP) [10]. In [10], the authors proposed an algorithm for extracting local directional texture patterns which are similar to the LDP transformation that we exploit. Both transformations work at a pixel level and use Kirsch masks. The image is split into small regions to compute a histogram of the features for each of them, and the feature vector is formed by concatenating these histograms. Finally, support vector machines are used for classification. Another method that exploits LBP was shown in [8], and we have adapted this algorithm in our research. A high-level outline is similar, but technical details are completely different (i.e., normalization, transformation, splitting and classification). Furthermore, our algorithm includes an additional step—masking with a face relevance map [2].

Another method is presented in [11] by Kimura and Yachida. They proposed to analyze the degree of emotion beside emotion recognition. Their method is based on the idea that the degree of facial expression, such a subtle change, can be extracted as a variation from expressionless face and recognition can be achieved by analyzing it. In [12], Otsuka and Ohya model the motion with a hidden Markov model in such a way that each state corresponds to the conditions of facial muscles, e.g., contracting, apex and relaxing. The probability assigned to each state is updated iteratively.

In 1999 Lowe introduced scale-invariant feature transform (SIFT) [13] which is used in object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking, individual identification of wildlife and match moving. In [14, 15] it was used for emotion recognition.

Some authors perform analysis from facial image sequences, but such works are less popular than works about analysis of static facial images. Essa and Pentland describe a computer vision system for observing facial motion by using an optimal estimation optical flow method coupled with geometric, physical and motion-based dynamic models describing the facial structure [16].

### **3. Proposed algorithm**

Our method is based on an observation that not all parts of a face are significant in face recognition. Moreover, the data excess can even worsen recognition results. It can happen that in particular learning group people with particular haircut represent the same emotion by coincidence. Such cases can lead to overfitting.

Feature extraction in our method is carried out based on local directional pattern introduced in 2010 by Jabid, Kabir and Chae [9].

### 3.1. Local directional pattern transformation

Input of LDP consist of 2 items: matrix P representing an original image:

$$P = \begin{bmatrix} p_{0,0} & p_{0,1} & p_{0,2} & \dots \\ p_{1,0} & p_{1,1} & p_{1,2} & \dots \\ p_{2,0} & p_{2,1} & p_{2,2} & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}, \quad (1)$$

and a parameter  $k$  which is an integer number between 1 and 8. Output of LDP is a matrix of integers. Size of this matrix is the same as the size of the input matrix,  $k$  represents the number of neighboring pixels taken into account during comparison with neighbors.

$$\begin{array}{cccc} \begin{bmatrix} -3 & -3 & +5 \\ -3 & 0 & +5 \\ -3 & -3 & +5 \end{bmatrix} & \begin{bmatrix} -3 & +5 & +5 \\ -3 & 0 & +5 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} +5 & +5 & +5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} +5 & +5 & -3 \\ +5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \\ K_0 & K_1 & K_2 & K_3 \\ \begin{bmatrix} +5 & -3 & -3 \\ +5 & 0 & -3 \\ +5 & -3 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ +5 & 0 & -3 \\ +5 & +5 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ +5 & +5 & +5 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & +5 \\ -3 & +5 & +5 \end{bmatrix} \\ K_4 & K_5 & K_6 & K_7 \end{array}$$

Fig. 1. Kirsch compass masks

Rys. 1. Maski kierunkowe Kirscha

LDP transformation is based on Kirsch masks, which are  $3 \times 3$  masks applied to edge detection. There are 8 masks, presented in Fig. 1, and each of them is associated with one of basic 8 cardinal and intercardinal directions ( $K_0$  – east,  $K_1$  – northeast,  $K_2$  – north,  $K_3$  – northwest,  $K_4$  – west,  $K_5$  – southwest,  $K_6$  – south,  $K_7$  – southeast).

We calculate 8 matrices  $M_0 - M_7$  based on the formula  $M_i = P * K_i$ , where  $*$  is the convolution operation. Size of the matrices  $M_0 - M_7$  is the same as the size of matrix P. Their construction is presented in Fig. 2.

$$\begin{array}{ccc} \begin{bmatrix} m_{0,0}^0 & m_{0,1}^0 & m_{0,2}^0 & \dots \\ m_{1,0}^0 & m_{1,1}^0 & m_{1,2}^0 & \dots \\ m_{2,0}^0 & m_{2,1}^0 & m_{2,2}^0 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} & \begin{bmatrix} m_{0,0}^1 & m_{0,1}^1 & m_{0,2}^1 & \dots \\ m_{1,0}^1 & m_{1,1}^1 & m_{1,2}^1 & \dots \\ m_{2,0}^1 & m_{2,1}^1 & m_{2,2}^1 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} & \begin{bmatrix} m_{0,0}^2 & m_{0,1}^2 & m_{0,2}^2 & \dots \\ m_{1,0}^2 & m_{1,1}^2 & m_{1,2}^2 & \dots \\ m_{2,0}^2 & m_{2,1}^2 & m_{2,2}^2 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \end{array}$$

Fig. 2. Matrices  $M_0 - M_7$

Rys. 2. Macierze  $M_0 - M_7$

Once we get the matrices  $M_0 - M_7$ , we calculate the final LDP value:

$$LDP_k(x, y) = \sum_{i=0}^7 2^i g_k(m_{x,y}^i, \{m_{x,y}^j : j = 0, \dots, 7\}), \quad (2)$$

$$g_k(x, S) = \begin{cases} 1 & \text{if } x \in \max_k \{S\} \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

$$\max_k \{S\} = \text{set of top - } k \text{ elements in set } S. \quad (4)$$

This is computed for every pixel in the input image represented by matrix  $P$ .

Sample images before and after the transformation are shown in Fig. 3. Edges in direction of  $K_0$  mask (east) are clearly visible, but from human perspective other information are not apparent due to specific encoding procedure.

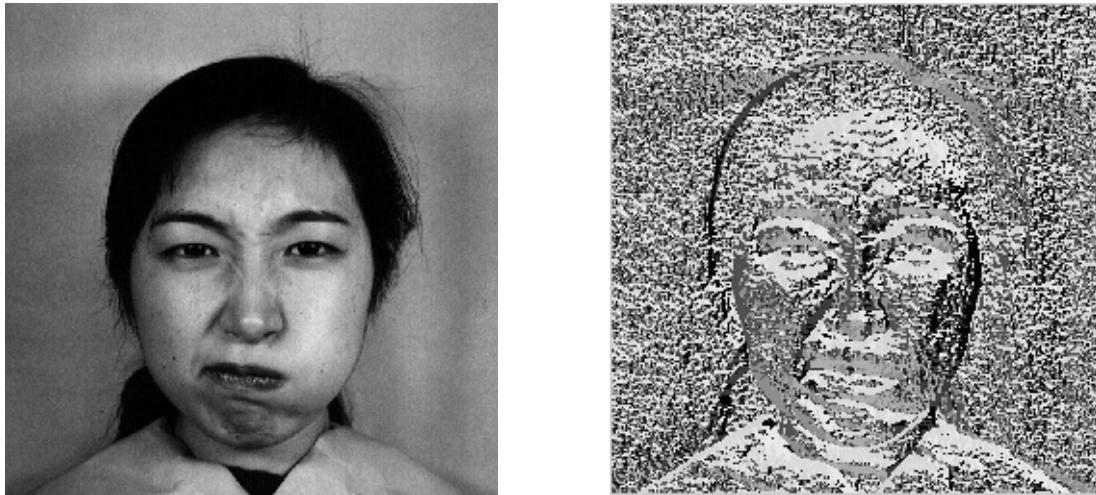


Fig. 3. Original image (left) and the image transformed using LDP (right)

Rys. 3. Oryginalne zdjęcie (z lewej) i zdjęcie po transformacie LDP (z prawej)

### 3.2. Splitting into subimages and masking using a face relevance map

We divide such a transformed image into some numbers of subimages of equal size which usually are of a square shape. The square side length is about half of the lips width.

After splitting into subimages, we mask the whole image using a binary face relevance map. Some subimages are taken for further analysis, while some are discarded. A few sample masks are provided in Fig. 4. The discarded subimages are not involved in the further recognition in any way.

### 3.3. Calculating histograms

For each subimage (which was not discarded during masking with the face relevance map) we calculate a separate histogram of LDP results. LDP formula returns a binary number based on nearest pixels values. This binary number represents position of  $k$  largest neighbors.

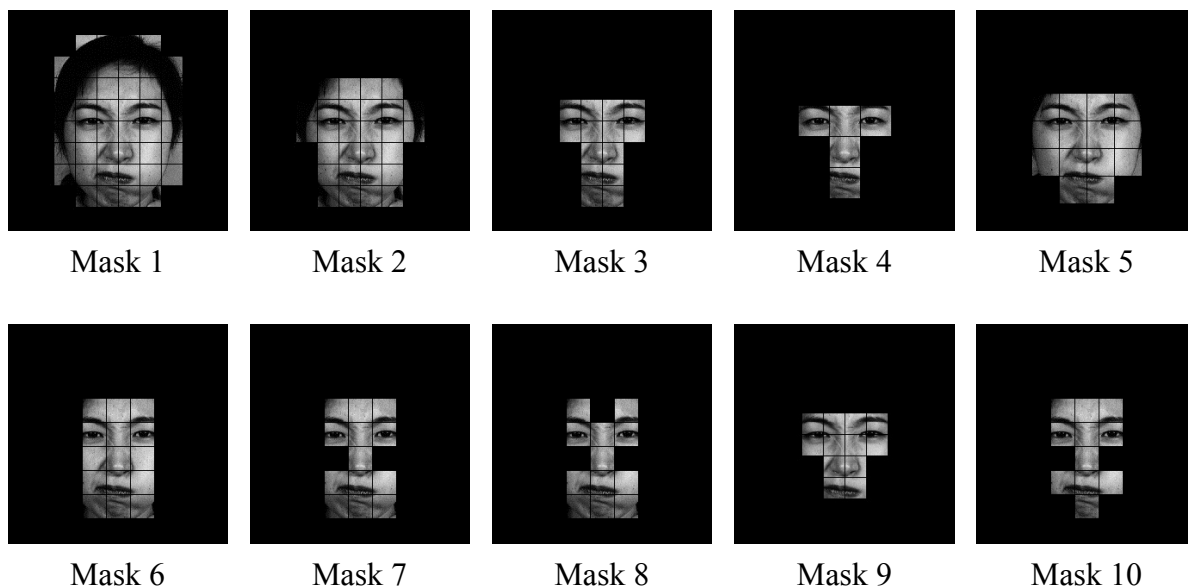


Fig. 4. Sample binary masks

Rys. 4. Przykładowe maski binarne

The returned values are within the range  $00000000_2$  ( $0_{10}$ ) to  $11111111_2$  ( $255_{10}$ ), but only a certain number of all 256 values are permissible. The returned binary value has a constant number of '1's, which equals  $k$ . If  $k$  is 3, then for example  $00000111_2$  ( $7_{10}$ ) or  $01000101_2$  ( $69_{10}$ ) are the permissible values, while  $01000111_2$  ( $71_{10}$ ) or  $01000100_2$  ( $68_{10}$ ) are not. It

means that we get  $\binom{8}{k}$  permissible values in the subimage histogram. The histogram lengths

for different  $k$  are shown in Table 1.

Table 1

Histogram lengths for different values of $k$	
$k$	Histogram length
0	1
1	8
2	28
3	56
4	70
5	56
6	28
7	8
8	1

After we get a histogram for each subimage, we concatenate all histograms into one long feature vector.

Once we get histograms for all images from learning collection, we check if some feature value equals zero for every image in the training set. In such a case, we delete that very feature from all the histograms. This additionally reduces the length of the histograms.

### 3.4. Classification

Support vector machines (SVMs) are supervised learning models with associated learning algorithms that can be used to analyze data and recognize patterns. SVMs are used to solve linearly separable classification problems. This is done by finding an optimal hyperplane which divides the classes. The hyperplane is determined by maximizing the margin which separates the data (this is the distance that the hyperplane could be moved before hitting a datapoint).

The original optimal hyperplane algorithm proposed by Vapnik and Lerner in 1963 [17] was a linear classifier. However in next years, Boser, Guyon and Vapnik suggested a way to create nonlinear classifiers by applying the kernel trick to maximum-margin hyperplanes [18]. Some common kernels are presented in Table 2 –  $u$  and  $v$  represent two different feature vectors,  $d$  is the polynomial degree,  $\sigma$  is the kernel width parameter,  $\kappa$  is the slope and  $c$  is the intercept constant.

Table 2

Kernels and theirs formulas

Kernel	Formula
Homogeneous polynomial	$K(u, v) = (u \cdot v)^d$
Inhomogeneous polynomial	$K(u, v) = (u \cdot v + 1)^d$
Gaussian radial basis function	$K(u, v) = e^{-\frac{ u-v ^2}{\sigma^2}}$
Hyperbolic tangent	$K(u, v) = \tanh(\kappa u \cdot v + c)$

Based on our experiments, we decided to use the radial basis function kernel, because it delivered the best results.

The basic SVM was designed for 2-classes problems, but a set of binary classifiers can be applied to multiclass classification. There are 3 possible approaches to build a multiclass classifier:

- using a base class,
- 1-to-N comparisons,
- 1-to-1 comparisons.

In the *base class* approach, one class is treated as a base class, and the remaining classes are classified against this one. The strongest response wins. It means that  $(N-1)$  2-class classifiers have to be learned. Main advantage of this method is the linear complexity, and the main disadvantage lies in poor ability of separating non-base classes.

In *1-to-N comparisons*, we compare each class against the rest. The strongest response wins. This approach is more universal because of lack of a base class, moreover complexity is still linear (as in the *base class* approach) so this method is also fast.

In *1-to-1 comparisons* approach, we compare every class against each other. This results in  $(N(N-1)/2)$  2-class classifications, therefore complexity is quadratic and this method is slower than others. The most important benefit of this approach lies in the best precision.

We decided to use the second approach, i.e. *1-to-N comparisons*, because it is a good compromise between the precision and execution time.

#### 4. Experimental validation

In our experiments, we have focused on checking if discarding subimages may improve the recognition quality. In a usual case ( $k$  equals 3, the image is divided into 100 subimages and about 3/4 of them are discarded) we get feature vector with the length of  $\binom{8}{3} \times (64 - 48) = 896$ . To perform the classification we use multiclass SVMs (also known as support vector networks) [19]. Here, each emotion represents a single class.

We used the images from the Japanese Female Facial Expressions (JAFFE) database. This database contains 213 digital images of 10 female persons, who manifest 7 basic emotions (anger, disgust, fear, happiness, neutrality, sadness and surprise). The images count of particular emotions and persons are roughly the same (see Table 3). The images in this database are in the grayscale, therefore we operate using a single channel only.

Table 3

	KA	KL	KM	KR	MK	NA	NM	TM	UY	YM	Sums
Anger	3	3	3	3	3	3	3	3	3	3	30
Disgust	3	4	2	3	3	3	2	3	3	3	29
Fear	4	3	3	3	3	3	3	3	3	4	32
Happiness	4	3	4	2	3	3	3	3	3	3	31
Nervous	3	3	3	3	3	3	3	3	3	3	30
Sadness	3	3	4	3	3	3	3	3	3	3	31
Suprise	3	3	3	3	3	3	3	3	3	3	30
Sums	23	22	22	20	21	21	20	21	21	22	213



JAFPE database contains 213 images, which is not much. Hence, we decided to use about 90% of the images for learning (our splitting procedure does not guarantee to always split in the exactly same proportion). For each mask and parameters combination, we perform 10-fold cross validation and count the average result and its standard deviation.

We have conducted tests for several masks with different resolutions and the number of subimages (see Table 4). It occurred that the smaller (with less subimages count) mask, the better result, but at some point deleting another subimages lead to worse accuracy of recognition. Compare mask no. 3 with 9, 7 with 8 or 7 and 10 and their average result of recognition.

In our opinion, our algorithm is very sensitive to the translations that cause a certain part of the face be positioned in a different subimage than usually. One way for handling this situation is to improve the image normalization (resizing, moving and rotation). Unfortunately, this solution sometimes is not sufficient, because not all people have faces with the same proportion.

Our best recognition rate (using mask no. 3) is 92.5% and it is better than reported in [20] and [21], but worse than in [8], which rely on much more advanced normalization techniques. It is worth noting that the mask type greatly influences the classification accuracy. One of the reasons that may affect the final score lies in some ambiguities in the ground-truth annotations. Some examples are presented in Fig. 5. *Original label* is the label assigned by the database authors.

Table 4

Results of experiments

Method	Mask	Dimension	Subimage count	Classification accuracy
Our	without mask	9×9	81	79.44% ± 10.9%
Our	without mask	10×10	100	76.41% ± 12.12%
Our	no. 1	10×10	50	86.88% ± 9.47%
Our	no. 2	10×10	30	87.43% ± 9.03%
Our	no. 3	10×10	14	92.46% ± 7.17%
Our	no. 4	7×7	5	87.05% ± 8.48%
Our	no. 5	8×8	14	87.49% ± 9.59%
Our	no. 6	9×9	15	91.12% ± 7.1%
Our	no. 7	9×9	13	91.71% ± 7.06%
Our	no. 8	9×9	12	91.64% ± 7.76%
Our	no. 9	10×10	12	89.83% ± 6.94%
Our	no. 10	9×9	11	87.52% ± 6.8%
Zhang et al. [20]				90.1%
Guo et al. [21]				91%
Feng et al. [8]				93.8%

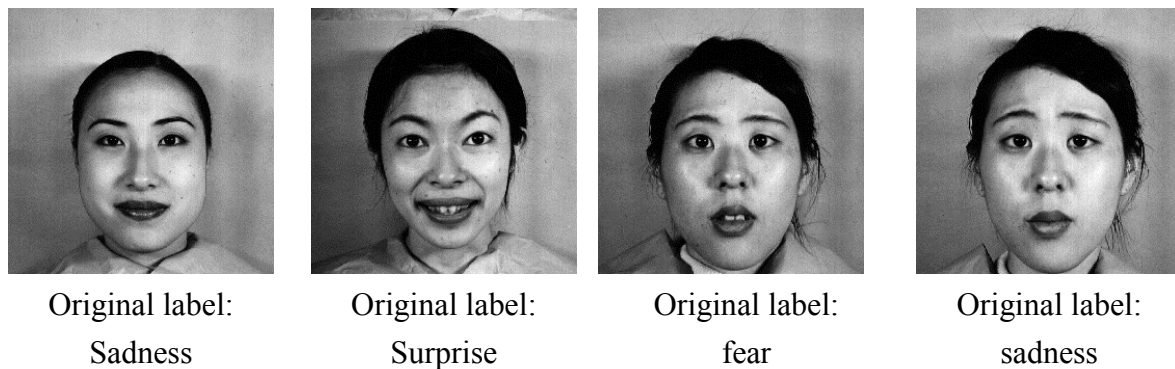


Fig. 5. Examples of questionable annotations  
Rys. 5. Przykłady błędnie oznaczonych fotografii

## 5. Conclusions and Future Work

This paper reports a study on emotion recognition from static grayscale images. The main contribution consists in using binary face relevance maps. In the reported experimental study, we demonstrated that the expression recognition accuracy for the JAFFE database can be substantially improved and it is better than some of the results reported in the literature.

In our future work, we want to improve the approach outlined in this paper. Our preliminary vision is to detect facial landmarks (e.g. eyes, nose, mouth, eyebrows), and use them to determine locations of the subimages instead of analyzing the same parts of the entire image. Moreover, we want to improve the face relevance maps: currently they are binary and manually created. We want them to be composed of continuous values and generated automatically. Face relevance maps optimization techniques have been proposed and reported successful for face recognition in our earlier study [2].

Moreover, we want to validate our approach using other databases, because JAFFE database is rather meager – it contains only 213 images, which is too small to perform proper SVM learning. Other databases with images representing basing emotions contain much more images, e.g. *Facial Expressions In The Wild* database contains 957 samples representing 6 expression classes and a neutral, *MMI Facial Expression Database* contains over 2900 videos and high-resolution still images of 75 subjects.

**BIBLIOGRAPHY**

1. Ekman P., Hager J. C., Oster H.: *Emotion in the Human Face*. Cambridge University Press, 1982.
2. Kawulok M., Wu J., Hancock E. R.: Supervised relevance maps for increasing the distinctiveness of facial images. *Pattern Recognition*, vol. 44, 2011, p. 929÷939.
3. He D.-C. and Wang L.: Texture unit, texture spectrum, and texture analysis. *Geoscience and Remote Sensing*, vol. 28, 1990, p. 509÷512.
4. Ojala T., Pietikainen M., Maenpaa T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence*, vol. 24, 2002, p. 971÷987.
5. Ahonen T., Hadid A., Pietikainen M.: Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence*, vol. 28, 2006, p. 2037÷2041.
6. Zhao G., Pietikainen M.: Dynamic texture recognition using local binary patterns with an application to facial expressions. *Pattern Analysis and Machine Intelligence*, vol. 29, 2007, p. 915÷928.
7. Feng X., Pietikainen M., Hadid A.: Facial expression recognition with local binary patterns and linear programming. *Pattern Recognition and Image Analysis*, vol. 15, 2005, p. 546÷548.
8. Shan C., Gong S., McOwan P. W.: Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vision Comput.*, vol. 27, 2009, p. 803÷816.
9. Jabid T., Kabir M., Chae O.: Local directional pattern (ldp) for face recognition. *Consumer Electronics*, January 2010, p. 329÷330.
10. Rivera A. R., Castillo J. R., Chae O.: Local directional texture pattern image descriptor. *Pattern Recognition Letters*, vol. 51, 2051, p. 94÷100.
11. Kimura S., Yachida, M.: Facial expression recognition and its degree estimation. *Computer Vision and Pattern Recognition*, June 1997, p. 295÷300.
12. Otsuka T., Ohya J.: Spotting segments displaying facial expression from image sequences using hmm. *Automatic Face and Gesture Recognition*, April 1998, p. 442÷447.
13. Lowe D.: Object recognition from local scale-invariant features. *Computer Vision*, vol. 2, 1999, p. 1150÷1157.
14. Soyel, H., Demirel H.: Improved sift matching for pose robust facial expression recognition. *Automatic Face Gesture Recognition and Workshops*, March 2011, p. 585÷590.

15. Tariq U., Lin K.-H., Li Z., Zhou X., Wang Z., Le V., Huang T., Lv X., Han, T.: Emotion recognition from an ensemble of features. Automatic Face Gesture Recognition and Workshops, March 2011, p. 872÷877.
16. Essa I., Pentland A.: Coding, analysis, interpretation, and recognition of facial expressions. Pattern Analysis and Machine Intelligence, vol. 19, 1997, p. 757÷763.
17. Vapnik V.: Pattern recognition using generalized portrait method. Automation and Remote Control, vol. 24, 1963, p. 774÷780.
18. Boser B. E., Guyon I. M., Vapnik V. N.: A training algorithm for optimal margin classifiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory, 1992, p. 144÷152.
19. Cortes C. and Vapnik V.: Support-vector networks. Machine Learning, vol. 20, 1995, p. 273÷297.
20. Zhang Z.: Feature-based facial expression recognition: Sensitivity analysis and experiments with a multilayer perceptron. International Journal of Pattern Recognition and Artificial Intelligence, vol. 13, 1999, p. 893÷911.
21. Guo G., Dyer C.: Simultaneous feature selection and classifier training via linear programming: a case study for face expression recognition. Computer Vision, June 2003, p. I-346÷I-352.

## Omówienie

Istnieje 6 podstawowych emocji, które są uniwersalne pośród wszystkich ludzi niezależnie od rasy i kultury: gniew, wstręt, strach, szczęście, smutek i zaskoczenie. Rozpoznawanie takich emocji na podstawie zdjęć jest obiektem zainteresowania psychologów, ale może również dostarczać interesujących informacji dla służb ochrony lotnisk czy stacji kolejowych. Emocje można rozpoznawać zarówno na podstawie pojedynczych zdjęć, jak i sekwencji wideo. Pierwsze podejście jest trudniejsze, ponieważ mamy mniej danych, czyli również mniej informacji, jednakże z reguły jest łatwiejsze w implementacji. W przypadku rozpoznawania emocji na podstawie pojedynczego zdjęcia szczególnie istotne jest posiadanie zdjęcia wykonanego w odpowiednim momencie. Kiedy ktoś zaczyna wyrażać jakąś emocję, to początkowo trudno jest ją rozpoznać. Jednakże po krótkiej chwili emocja jest o wiele wyraźniejsza, a potem znowu staje się mniej charakterystyczna. W artykule tym zaprezentowano nowe podejście do kwestii rozpoznawania emocji. Najpierw wykonywaliśmy transformację LDP, następnie dzieliliśmy obrazek na regiony, na które była nakładana binarna mapa istotności twarzy (ang. *face*

*relevance map*). Na koniec otrzymany wektor cech został użyty do uczenia nadzorowanego z wykorzystaniem maszyny wektorów podpierających. W sekcji 4 pokazano wpływ doboru maski na dokładność rozpoznawania emocji oraz porównano otrzymane wyniki z wynikami uzyskanymi przez innych badaczy.

### Addresses

Tomasz HERUD: Silesian University of Technology, Institute of Informatics, ul. Akademicka 16, 44-100 Gliwice, Poland, [tomasz.herud@polsl.pl](mailto:tomasz.herud@polsl.pl). Future Processing, ul. Bojkowska 37A, 44-100 Gliwice, Poland, [therud@future-processing.com](mailto:therud@future-processing.com).

Michał KAWULOK: Silesian University of Technology, Institute of Informatics, ul. Akademicka 16, 44-100 Gliwice, Poland, [michal.kawulok@polsl.pl](mailto:michal.kawulok@polsl.pl). Future Processing, ul. Bojkowska 37A, 44-100 Gliwice, Poland, [mkawulok@future-processing.com](mailto:mkawulok@future-processing.com).

Bogdan SMOLKA: Silesian University of Technology, Institute of Automatic Control, ul. Akademicka 16, 44-100 Gliwice, Poland, [bogdan.smolka@polsl.pl](mailto:bogdan.smolka@polsl.pl).