

*Basic emotions, recognition, intrinsic dimensionality, geometrical wavelets*

Agnieszka LISOWSKA<sup>\*</sup>, Wiesław KOTARSKI<sup>\*</sup>,  
Aron KOWNACKI<sup>\*</sup>

## **SEMI-AUTOMATIC HUMAN EMOTIONS RECOGNITION METHOD BASED ON I2D FEATURES**

In the paper we present quite new approach to the problem of human emotion recognition with use of face images. We assume that basic emotions such as anger, disgust, fear, happiness, sadness, surprise are expressed by face mimic. Face images with the well defined emotions may be performed using the method based on geometrical wavelets (beamlets) in order to extract intrinsically two dimensional features, the most important ones from the Human Visual System point of view. Such an approach can be successfully applied in extraction process of the most important features that are responsible for recognition of basic elements of face (eyes, nose, lips, etc.). The listed elements of face have a little different location that depends on emotion expressed. It has been proved experimentally that it is possible using very small amount of information extracted from a face image, by the so-called beamlet extractor, to recognize emotion with high accuracy. Very promising results of experiments suggest that the method should be further investigated and improved.

### **1. INTRODUCTION**

Investigations in automatic face recognition, identification and also emotion recognition systems based on face images are of great interest in practical applications [4], [12], [14], [15]. Emotion recognition systems might be of some interest not only for medicine, especially for psychology and psychiatry, but also in social sciences.

Encouraged by the promising results concerning the recognition of human faces with the use of intrinsically two dimensional features, reported in [10], we posed the question if is it possible to recognize emotions presented in face images. Our starting point was the collection of faces images that have expressed different well defined emotions. Images have been analysed with the help of the application EmoReco [7] prepared by the third author.

We would like to recall some introductory information that stands behind the idea of emotion recognition method and creates theoretical background of the method presented in this paper.

Recent investigations in neuropsychology and psychology of vision have approved that human eye filters information from the surrounding world in a quite specific way. Indeed, information received by human brain may be classified into three different kinds of signals [6]. Shortly speaking, the less important ones, at first look, are textures and the most important ones are different types of corners, junctions, end of lines, etc. Between these two types are straight lines and edges. As a consequence of those facts the notion of intrinsic dimensionality used in image processing has arisen [17]. Also the wide spectrum of different classes of feature extractors have been constructed [13]. One extractor has been proposed by the first author and described in [8], [9]. That extractor has been used as the main tool in the process of emotion recognition.

---

<sup>\*</sup> Institute of Informatics, Silesian University, Będzińska 39, 41-200 Sosnowiec, Poland

It is known that every face is recognised by human eye thanks to the presence of basic elements in it such as eyes, nose, lips, etc. [2], [11]. To a high degree just these features determine the face image. It is because of the fact that the main important information (junctions, corners, etc.) is localised nearby basic elements of the face. Obviously, the listed above elements of face have a little different location that depends on emotion expressed. The most important face's features can be extracted automatically from a face image by the beamlet extractor [10].

Experimental results, described in the paper, have approved that beamlets based feature extractor can get out from an image the information related to face's elements according to psychology of vision. Basing on that very small amount of information it is possible to recognise properly the emotion expressed by the face image. Obtained results of experiments showed high accuracy in recognition of emotion expressed in the face images.

## 2. BASIC EMOTIONS EXPRESSED BY FACE MIMICS

Popular meaning of emotions is obvious for everybody. Really, if we ask anybody "Do you know what emotion is?", we get the answer "yes". It supprises nobody, because emotions accompany people always. We can even recognize if somebody is sad or happy. Emotions are usually expressed by human face. But from psychological point of view intuitive emotions understanding is not enough. That is because it is not possible to answer for the most important questions e.g.:

- What differs emotion, mood, feeling and temperament?
- How to classify emotions?

Now one can see that the answers to the posed above questions are not so easy.

There are many psychological theories which try to find the precise definition of emotion and their adequate classification. But we do not want to go into details. For our purposes it will be enough to use Ekman's basic emotions classification [2]. He distinguished the following basic emotions: anger, disgust, fear, happiness, sadness and suprise.

In our further considerations we assume that we have a collection of face images representing properly classified basic emotions. Basing on that collection, using image processing techniques, we show that it is possible to extract features that can adequately characterize a given emotion. Applied image processing techniques are strongly related to properties of the so-called Human Visual System (HVS), the idea of intrinsic dimensionality and a quite new tool – geometrical wavelets (beamlets). More details will be presented in the next section.

## 3. PERCEPTION OF IMAGES

In the following two subsections some important properties of HVS and the notion of intrinsic dimensionality will be described. It will be explained how the visual information is perceived by a man, what kind of information is the most important to properly understand contents of an image and what theoretical explanation stands behind this.

### 3.1. PROPERTIES OF HUMAN VISUAL SYSTEM

From experiments it is known that human eye observing an image must process a great amount of information. But due to the fact that human memory is limited, only a small piece of information from the eye is sent to the brain. It means that information about an image must be filtered – one part of it goes to the brain, while the rest is stopped and ignored [3]. Many

psychovisual experiments have shown that there exist classes of information which human eye perceives on different levels of significance.

It was noticed that curvature points and corners, junctions, etc. bring more information about image than straight lines or edges due to the fact that they cannot easily be predicted from neighbouring points [5]. It was showed also that some simple objects might be reconstructed from images containing only curvature points and corners by simply connecting them in an appropriate way.

Further experiments reported in the literature proved the existence of curvature selective cells that make the human eye sensitive to curvature points, junctions, corners, etc. Human eye uses information from orientation selective cells. Therefore, human eye is also sensitive to directions present in images.

As follows from the above considerations there is a special kind of information, which is perceived by human eye in the first order. It is believed that this information, that is junctions, corners, ends of lines, etc. make up the smallest possible amount of information about an object, which human brain needs to reproduce the object in memory.

Theoretical explanation of the facts described above has been found. Namely, using the notion of intrinsic dimensionality it is possible to understand how HVS works. Also the theory of features extractors (based on intrinsic dimensionality) that can automatically extract prescribed features from digital images, in the same way as human eye can do it, has appeared.

### 3.2. INTRINSIC DIMENSIONALITY

The notion of intrinsic dimensionality was first introduced to image processing by Zetzsche and Barth [17]. It was introduced to differentiate mathematically the local and global dimensions present in images. But it also well reflects the degrees of their importance in image perception by human eye. In one of the commonly used definitions of intrinsic dimensionality the three kinds of regions may be found in an image. Namely,

**i0D** – textures,

**i1D** – straight lines and edges,

**i2D** – otherwise – it means the whole rest (junctions, corners, ends of lines, etc.).

This definition divides an image into three kinds of disjoint (or sometimes with slightly fuzzy boundaries) areas, respectively – smooth parts of an image, straight edges and edges, which are not straight lines. It is worth mentioning also that all three kinds of intrinsic dimensionality occur in images with different probability [13]. The more probable ones are i0D areas, while the less probable ones are i2D areas. Indeed, the great majority in images make up textures. However corners and junctions are in minority. Paradoxically, the last ones are the most important in image perception. Thus extraction of them from images is very important task because it allows bringing out the main information from an image.

There are many different approaches to feature extraction. Eg. in [10] one can find more details about them. We will use the one based on beamlets proposed in [9].

### 3.3. I2D FEATURE EXTRACTOR BASED ON BEAMLETS

Before we present the definition of the feature extractor based on beamlets let us introduce some basic facts concerning theory of beamlets [1]. Define an image domain as a square  $S = [0,1] \times [0,1]$ . Let us assume that in the image smooth edges are present. Consider then the dyadic square  $S(k_1, k_2, j)$  as a collection of points such that

$$S(k_1, k_2, j) = [k_1 / 2^j, k_1 + 1 / 2^j] \times [k_2 / 2^j, k_2 + 1 / 2^j],$$

where  $0 \leq k_1, k_2 < 2^j$  for integer  $j \geq 0$ . Note that  $S(0,0,0)$  denotes the whole image domain, that is the square  $[0,1] \times [0,1]$ . On the other hand  $S(k_1, k_2, j)$  for  $0 \leq k_1, k_2 < N$  denote appropriate pixels from  $N \times N$  grid, where  $N$  is dyadic ( $N = 2^j$ ).

Having assumed that an image is defined on the square  $[0,1] \times [0,1]$  and that it consists of  $N \times N$  pixels, one can note that on each border of any square  $S(k_1, k_2, j)$ ,  $0 \leq k_1, k_2 < 2^j$  there are vertices with distance equal to  $1/N$ . Every two such vertices in any fixed square may be connected to form a straight line called *beamlet* [1]. The set  $B$  of all possible beamlets within all possible dyadic squares we call the *beamlets' dictionary*. Let us notice that it consists of beamlets, which differ in all possible locations, scales and orientations. Thanks to it such representation allows to determine any edge present in image with quite good exactness.

In [8] the following definition of feature extractor called the *i2D Selective Beamlet Operator* (i2D SBO) has been proposed

$$F_B(x_1, x_2) = \sum_{J, k_1, k_2, m} \alpha_{J, k_1, k_2, m} b_{J, k_1, k_2, m}(x_1, x_2),$$

where  $J = \lg N$ ,  $0 \leq k_1, k_2 < 2^j$ ,  $0 \leq m \leq M(S(k_1, k_2, j))$  and  $\alpha_{J, k_1, k_2, m} \in \{0, 1\}$ ,  $b_{j, k_1, k_2, m} \in B$ .

The definition of the operator is based on the fact that from the intrinsic dimensionality point of view all straight fragments of edges (the i1D ones) are approximated by beamlets lying within squares of size larger than one. All other fragments that are not straight (that are corners, junctions and ends of lines – the i2D ones) lie in squares that are simple pixels (indeed they may not be approximated by any longer beamlet).

#### 4. EXPERIMENTAL RESULTS

Face images used in experiments have been taken with the help of a standard digital camera in almost the same conditions (illumination, distance). The obtained images have been converted first to binary bitmaps of  $128 \times 128$  pixels in resolution using Corel Photo Paint software. Then, EmoReco application has been used. EmoReco works in the following way: it extracts i2D features from every image as it is demonstrated in Fig. 1.

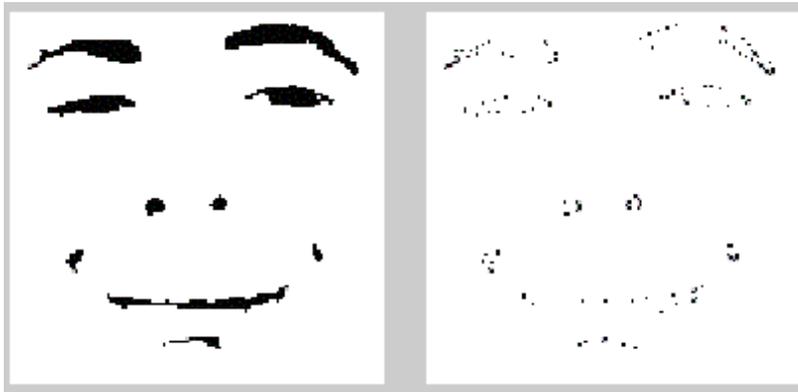


Fig. 1. Image (left) and extracted i2D features (right).

Next, images which present only i2D features are compared basing on intuitively constructed normalized similarity measure  $\varphi: d_{av} \rightarrow [0, 1]$  defined according to the following formula:

$$\varphi(d_{av}) = \begin{cases} 0, & d_{av} > 25 \\ 1 - \frac{d_{av}}{25}, & d_{av} \leq 25 \end{cases},$$

where  $d_{av} \in [0, 255]$  represents the average displacement between  $N$  pixels of compared ( C ) and  $M$  pixels of pattern ( P ) images calculated according to the formula:

$$d_{av} = \frac{\sum_{i=1}^N \min\{d(i, j), j \in \{1, \dots, M\}\}}{N}, \quad d(i, j) = |x_{iC} - x_{jP}| + |y_{iC} - y_{jP}|,$$

where  $x_{iC}$ ,  $y_{iC}$ ,  $x_{jP}$ ,  $y_{jP}$  are coordinates of pixels in compared ( C ) and pattern ( P ) images, respectively.

Two types of experiments have been performed. Experiment 1 has been carried out in the following way. The image-base containing patterns (shown in Fig. 2) of six basic emotions (according to Ekman [2]) has been created for one person. Next, 13 different face images (shown in Fig. 3) of the same person have been compared with the help of EmoReco software. The results of comparison are presented in Table 1.



Fig. 2. Emotion’s patterns. From left to right: disgust, happiness, sadness, fear, surprise and anger, respectively.

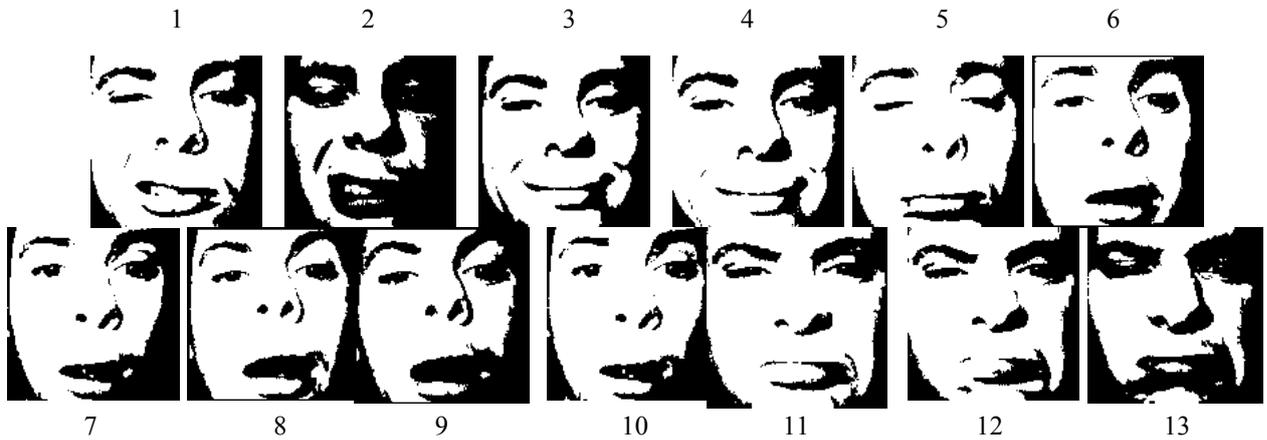


Fig. 3. Tested images in Experiment 1. Face’s images and patterns belong to the same person.

From Table 1 it is clearly seen that the highest similarity coefficient (marked in bold) show the right emotion in the most cases. However, in one case we obtained the same value of similarity measure for “fear” and “surprise” what means that the both emotions have not been exactly separated and one mistake had happened. Despite of that, the results obtained are satisfactory.

Table 1. Similarity coefficient between tested and pattern images. Pattern and tested images concern the same person.

Img.	disgust	happiness	Sadness	fear	suprise	anger
1	<b>0,912</b>	0,829	0,807	0,809	0,827	0,816
2	<b>0,735</b>	0,717	0,554	0,666	0,624	0,703
3	0,821	<b>0,913</b>	0,713	0,755	0,757	0,825
4	0,850	<b>0,872</b>	0,791	0,791	0,771	0,447
5	0,946	0,819	<b>0,914</b>	0,826	0,860	0,914
6	0,858	0,808	0,864	<b>0,899</b>	0,881	0,799
7	0,831	0,788	0,806	<b>0,868</b>	0,850	0,820
8	0,846	0,828	0,821	0,851	<b>0,914</b>	0,804
9	0,821	0,800	0,761	<b>0,856</b>	0,836	0,817
10	0,834	0,909	0,803	<b>0,852</b>	<b>0,852</b>	0,802
11	0,833	0,853	0,796	0,776	0,806	<b>0,862</b>
12	0,831	0,831	0,805	0,799	0,817	<b>0,874</b>
13	0,780	0,794	0,696	0,763	0,748	<b>0,857</b>

Encouraged and inspired by Experiment 1 we performed Experiment 2 results of which might be more risky and difficult for prediction. Namely, we have compared 5 face images of one person and 2 of another one (shown in Fig. 4) with the image-base of patterns obtained in Experiment 1. That means that compared images and patterns belonged to different persons.

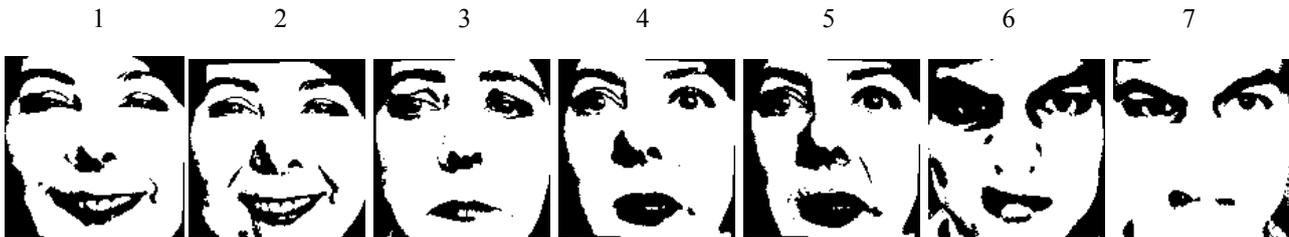


Fig. 4. Tested images in Experiment 2. Images 1- 5 present one person’s face, while 6-7 belong to the second person.

In Table 2 the values of similarity coefficient are gathered. In bold are marked the highest values, while in bold italic those expected. The results of comparisons were satisfactory for the first person (images 1-5), especially for face images expressing happiness and sadness. In the last two cases for the first person (images 4 and 5) mistakes had happened. Unfortunately, for the second person (images 6 and 7) we have obtained unsatisfactory results.

Table 2. Similarity coefficient between tested and pattern images. Pattern and tested images concern the different persons.

Img.	disgust	happiness	sadness	fear	suprise	Anger
1	0,760	<b>0,787</b>	0,714	0,726	0,739	0,774
2	0,732	<b>0,774</b>	0,685	0,691	0,698	0,758
3	0,720	0,701	<b>0,742</b>	0,705	0,718	0,731
4	0,723	0,754	0,705	0,714	<b>0,732</b>	<b>0,764</b>
5	0,715	0,744	0,680	0,706	<b>0,722</b>	<b>0,760</b>
6	<b>0,715</b>	<b>0,747</b>	0,692	0,701	0,710	0,740
7	<b>0,637</b>	0,671	<b>0,713</b>	0,641	0,668	0,694

The source of mistakes comes from, as we think, binarisation process during that some essential information might be lost or bad emotion's pattern used, not suitable for the second person.

## 5. CONCLUSIONS

The results of carried out experiments are promising. They proved that EmoReco application works very well and very good simulates behaviour of human brain during emotion recognition using face images. It is unbelievable that from a very small amount of information recognition process is possible. Only about in average 2% of pixels from the whole image in total are enough to obtain characterization of individual features of the face that are responsible for emotion expression. That fact is very important from database point of view.

It is also worth mentioning that the theory of beamlets is not complicated from mathematical point of view. It means that the extractor based on geometric wavelets is characterised by comparatively high speed in extraction of features from images.

The method presented in the paper is called as semi-automatic. That is why, face images further used in emotion's recognition process must be first manually preprocessed in order to obtain square binary images with face exactly in the centre. The authors think that it is possible to work out fully automatic system for human emotion recognition and identification basing on the approach presented in the paper. But it requires further investigations. Especially, it should be prepared larger database of images or investigations should be carried out on images that are available from [16]. Then, the obtained results might be statistically approved.

Further directions of investigations that are closely related to medicine are also possible to establish. One of the possibilities appears in the situation of evaluation of patient pain's intensity. Really, it is a well known fact that there is a high correlation between the level of pain occurring during many diseases and the appearance of the patient's face.

It seems to be possible to combine the method described in the paper with that based on neural networks presented in details in [15].

## BIBLIOGRAPHY

- [1] DONOHO D. L., HUO X., Beamlet Pyramids: A New Form of Multiresolution Analysis, Suited for Extracting Lines, Curves and Objects from Very Noisy Image Data, Proceedings of SPIE, Vol. 4119, 2000.
- [2] EKMAN P., FRIESEN W. V., Emotion in the human face: Guidelines for research and an integration of findings, New York, Pergamon Press, 1972.
- [3] GABOR D., Guest Editorial, IRE Transactions on Information Theory, September, 1959.
- [4] JIA X., NIXON M., S., Extending the Feature Vector for Automatic Face Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 12, December 1995.
- [5] KAISER P. K., The Joy of Visual Perception, <http://www.yorku.ca/eye/thejoy.htm> .
- [6] KRIEGER G., ZETZSCHE C., Nonlinear Image Operators for the Evaluation of Local Intrinsic Dimensionality, IEEE Transactions on Image Processing, Special Issue on Nonlinear Image Processing, Vol. 5, No. 6, pp. 1026-1042, 1996.
- [7] KOWNACKI A., Automatic Recognition of Emotions Basing On face's Image, BSc Thesis, University of Silesia, Institute of Informatics, Sosnowiec, (in Polish), 2004.
- [8] LISOWSKA A., Intrinsic Dimensional Selective Operator Based On Geomerical Wavelets, Journal of Applied Computer Science, Vol. 12, NO. 2, 2004.
- [9] LISOWSKA A., Geometrical Wavelets and their Generalizations in Digital Image Coding and Processing, PhD Thesis, University of Silesia, Institute of Informatics, Sosnowiec, 2005.
- [10] LISOWSKA A., KOTARSKI W., Automatic Perception of Significant Image Features Based on Psychology of Vision, Journal of Medical Informatics & Technologies, Vol. 7, pp. 31-40, 2004.

- [11] MAKIEŁA M., Methods of Faces' Image Identification Based on Characteristic Feature Points, MSc Thesis, Institute of Informatics, University of Silesia, Sosnowiec, (in Polish), 2002.
- [12] MAKIEŁA M., KOTARSKI W, LISOWSKA A., Automatic Human Face Recognition Method Based on Geometrical Face Features, Journal of Medical Informatics & Technologies, Vol. 5, pp. MI-47 – MI-56, 2003.
- [13] MITRA S. K., SICURANZA G. L., Nonlinear Image Processing, Academic Press, San Diego, 2001.
- [14] PAŁECZKA D., The Artificial Empathy. The Automatic Recognition Based on Facial Expressions, MSc Thesis, Poznań University of Technology, Poznań 2000, <http://www.cs.put.poznan.pl/dobek/empathy/>.
- [15] SHIN Y., Facial Expression Recognition Based Dimension Model Using Sparse Coding, Lecture Notes in Computer Science, Springer-Verlag, Vol. 3039, pp. 205-212, 2004.
- [16] THE DATABASE OF FACES, <http://www.uk.research.att.com/facedatabase.html> .
- [17] ZETZSCHE C., BARTH E., Fundamental Limits of Linear Filters in the Visual Processing of Two-Dimensional Signals, Vision Research, Vol. 30, pp. 1111-1117, 1990.