automatic perception,
intrinsic dimensionality

Agnieszka LISOWSKA*, Wiesław KOTARSKI**

AUTOMATIC PERCEPTION OF SIGNIFICANT IMAGE FEATURES BASED ON PSYCHOLOGY OF VISION

Recent investigations in neuropsychology and psychology of vision have proven that human eye does not get all the information from the surrounding world in the same degree. There are three classes of signals received by human brain. The more important one is the information about features such as corners, junctions, ends of lines, etc. Straight lines and edges are the second in the hierarchy of importance. And the last ones are textures they support the less important information about objects. Basing on these results, in image processing, theory of intrinsic dimensionality and related to it theory of feature extractors have been established. In the paper a survey of approaches that are used for construction of feature extractors based on intrinsic dimensionality have been presented. To carry out experiments the approach based on geometrical wavelets has been chosen and the software prepared by the first author has been used. Experiments presented in the paper have been performed on relatively complex images that had been faces' images. They confirmed that the information about the basic elements of faces (eyes, nose, lips, etc.) might be properly extracted from the face with the usage of the feature extractor. Moreover, the experiments have shown that in this way one could obtain the smallest possible amount of information, which was enough that human eyes yet have seen the face. Very promising results of experiments suggest that it is possible to use the proposed approach to face identification and recognition. Also some possible medical applications have been suggested.

1. INTRODUCTION

Recent investigations in neuropsychology and psychology of vision have confirmed that human eye filters information from the surrounding world in a quite specific way. Indeed, the information received by human brain may be classified into three different kinds of signals [4], [5]. 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 [23] has arisen. Also the wide spectrum of different classes of feature extractors has been constructed. All of them can extract the most important information from an image, but they differ in such aspects as effectiveness, time of working, or the possibility of some parameters changing, which allow for adjusting extractors to obtain the best possible results of their work.

Recently there is often a common need to have automatic systems to face identification and recognition [19]. Every face is recognised by human eye thanks to the

* Institute of Mathematics, Silesian University, Bankowa 14, 40-007 Katowice, Poland
** Institute of Informatics, Silesian University, Będzińska 39, 41-200 Sosnowiec, Poland

presence of basic elements in it such as eyes, nose, lips, etc. [12]. To a high degree just these features determine the face's image [18]. It is because of the fact that the main important information (junctions, corners, etc.) is localised nearby such basic elements of the face. Thanks to the wide spectrum of feature extractors such information may be automatically extracted from any image.

Experimental results, described in the paper, have confirmed that feature extractors can get out from an image the information related to face's elements according to psychology of vision. Moreover the experiments show that the extracted features form the smallest possible amount of information, basing on which, one can properly recognize the face.

2. PSYCHOLOGY OF VISION

Every day human eye must process very much information from surrounding world. But due to the fact that human memory is limited, only small piece of information from eye is sent to the brain. Indeed, in the retina there is some kind of filter, which passes some information and stops the other one [20]. But what kind of information is passed and what kinds are stopped? Many psychovisual experiments performed for about half a century have shown that there exist classes of information which human eye percepts on different levels of significance.

Already in 1954 Attneave [1] has 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. He also showed that some simple objects might be reconstructed from images containing only curvature points and corners by simply connecting them in an appropriate way. This approach is justified by some possible illusions. The one of them, and probably the most known, is presented in Figure 1. It shows the Kanizsa's figure [13]. It depicts four fragments of circles that are situated in such a way, that one can see the square which, in fact, is not present in the image. So one can "see" the object, probable even with edges, even if there is presented only some information on its corners. In reality there are no edges and textures present in image, which could define the square.

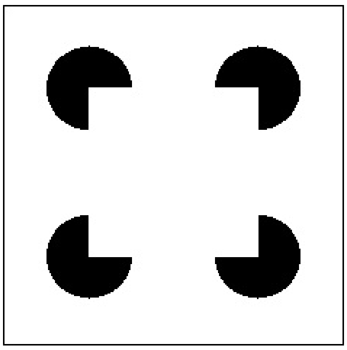

Fig. 1. Kanizsa's figure.

Besides illusions, it occurs the phenomenon of the cell adaptation and related to it the so-called after-effects phenomenon. A human eye has millions of light-sensitive receptors that are excited by light. While looking at an image the information about light obtained by receptors is sent to the brain and after that they take a moment to calm down. The observed image may remain in human mind for a moment. Such image is just after-image. It has been noticed the existence of the so-called curvature after-effects [7]. From experiments reported in the literature follows the existence of curvature selective cells. It means that the human eye is sensitive to curvature points, junctions, corners, etc. Human eye uses information from orientation selective cells. Indeed, human eye is also sensitive to directions present in images.

Many experiments confirmed that an object might be recognized basing only on edges. This means that textures are not so important and contain redundant information about image. It has been noticed that one can go even one step further, that is corners and junctions bring more information about image than straight lines or edges [4], [5].

Further experiments have shown that many cells present in eye are more sensitive to high curvature features than to the other ones [11]. Moreover they have shown also that within the high curvature selective cells one can do the more precise differentiation – different cells are sensitive to different kinds of curvature, as for example ends of lines, junctions, corners, etc. Great example of natural nonlinear filtering of high curvature points in eye concerns the frog’s retina. The detectors present in it are selective only to isolated dark spots on a brighter background. Indeed, none of them reacts to straight edges, lines, etc. Instead of the detectors respond to spots, line ends, curved edges and similar ones. Moreover, as it turned out, the selective cells are invariant to the location of the features. Indeed, it has been noticed that the human brain obtains the information separately about the location and the shape of the observed objects present in an image [14].

As follows from the above considerations there is a special kind of information, which is percepted 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.

3. AUTOMATIC FEATURE EXTRACTION

Basing on inventions from neuropsychology and psychology of vision the theory of intrinsic dimension used in computer graphics has arisen [23]. It concerns not only the notion of intrinsic dimensionality alone but also the theory of extraction of features of different intrinsic dimension from digital images in the same way as human eye can do it but in a quite automatic way. A great variety of different approaches concerning the problem of filtering of different features from digital images have appeared.

3.1. INTRINSIC DIMENSIONALITY

The notion of intrinsic dimensionality was first introduced to image processing by Zetzsche and Barth [23]. Though in reality this one, taken from statistics, is more general. It was introduced to differentiate the local and global dimensions present in images from mathematical point of view. But it also well reflects the degrees of importance in image perception by human eye. One of the two commonly used definitions of intrinsic dimensionality is based on the image's spectrum [15]. The three kinds of intrinsic dimensionality in image may be divided in the following way

i0D - the local spectrum is concentrated in the origin, the neighbour is constant (such is in the case of textures),

i1D - the local spectrum is concentrated in a line through the origin, it changes only in one direction (in the case of straight lines and edges),

i2D - otherwise – the local spectrum changes in many directions (in the case of junctions, corners, ends of lines, etc.).

Such definition divides an image into three kinds of disjoint (or sometimes with slightly fuzzy boundaries) areas, respectively – smooth parts of image, straight edges and edges, which are not straight lines. Recently also the continuous definition of intrinsic dimensionality has been introduced [16], but in common use is still the discrete one.

It is worth to mention also that all three kinds of intrinsic dimensionality occur in images with different probability [20]. 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.

3.2. DIFFERENT APPROACHES TO FEATURE EXTRACTION

There are many feature extractors in common use, which allow filtering digital images to get i2D features from them in automatic way. They are defined on the base of many different theories. Each approach has its advantages as well as disadvantages. Nevertheless, all extractors give more or less comparable results. Below there is given their classification according to the used approach.

1. Differential geometry. In nearly all practical approaches to the problem of detection of i2D signals the general concept that an image is the surface in three dimensional space is used. So the concept of intrinsic dimensionality may be related to the notion of the Gaussian curvature used in differential geometry. Indeed, planar regions of a surface correspond to i0D signals, parabolic regions to i1D ones and elliptic-hyperbolic to i2D ones [2].

In spite of the popularity of Gaussian curvature in differential geometry, in the case of i2D feature detection also other methods are used successfully. The second commonly used detector is based on the determinant of the so-called Hessian. The usage of the determinant of the Hessian as the i2D feature detector was introduced by Beaudet [3] and makes up the first constructed and quite effective corner detector.

2. Structure tensor. The idea of the structure tensor is that it approximates the autocovariance function in the origin. In all detectors based on structure tensor approach the eigenvalues as the main components are used. Depending on the rank of tensor one can determine the kind of intrinsic dimension. In the case of analysing smooth signals, i0D ones, all two eigenvalues are zero (but in practical use it assumes that they near zero). In the case of i1D signals one of the eigenvalues is large while the second one is zero (the rank of tensor equals to one). Whereas in the case of i2D features both of them are large (the rank of tensor is two). Hence the knowledge about the eigenvalues allows exactly determining the kind of a signal.

The one of the best known i2D detectors, based on the structure tensor approach, was introduced independently by Bigün and Granlund [6] and Förstner and Gülch [10].

3. Volterra series. Another of the main concepts related to the intrinsic dimensionality filtering theory is based on Volterra series [20]. Two dimensional theory of Volterra series has gained more popularity in recent years and is used in many areas of nonlinear image processing. One can see this theory as a generalization of Fourier series. The difference between the theories of Fourier and Volterra series is that in the former one only the linear part of series is taken into consideration, while the last one includes also the nonlinear parts (quadratic, cubic, etc.). Thanks to that, with the help of Volterra series (opposite to Fourier one), one can catch changes of signal in two directions simultaneously. So the signal is i2D if it changes at least in two directions, i1D one if changes only in one direction and i0D one otherwise.

Basing on all the above considerations Krieger and Zetzsche has proved that, with under some assumptions, the bilinear part of Volterra series may be used as an i2D feature extractor [15].

4. Geometrical wavelets. In recent days the theory of geometrical wavelets has been appeared [8], [9]. It makes up some kind of improvement of the classical theory of wavelets. Due to the fact that adaptive geometrical wavelets, such as beamlets or wedgelets, can extract from images straight lines or edges in different localization, scale and orientation they may be helpful in filtering i2D features in images. Indeed, all extracted well defined beamlets determine i1D signal, whereas degenerated one point ones determine i2D signal.

In [17] the extractor based on the adaptive geometrical wavelet approach has been defined.

5. Others. Besides the main approaches described above there are few others, which are applied in feature extraction. One of the known approaches is based on quadrature filters. Such approach allows estimating the local amplitude and phase of the signal. There are also other extractors, which are not based on any of the mentioned approaches, such as for example commonly used SUSAN extractor [21] or others.

There is a number of feature extractors, which are not based of any of the mentioned above approaches. But the ones described in this section give the main contribution to the theory of automatic feature extraction.

3.3. I2D FEATURE EXTRACTOR BASED ON BEAMLETS

As one can see there is a great variety of different feature extractors based on different approaches. The extractor used in performed experiments is based on geometrical wavelets, especially beamlets.

Before we present the definition of the feature extractor based on beamlets let us introduce some basic facts concerning theory of beamlets. Define an image 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

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

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

Having assumed that an image is defined within 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 $S(k_1, k_2, j)$, $0 \le k_1, k_2 < 2^j$ 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* [9]. 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 localizations, scales and orientations. Thanks to it such representation allows to determine any edge present in image with quite good exactness [9].

In the discrete case, with the above symbols, the Digital Beamlet Transform may be defined in the following way.

$$F(x_1, x_2) = \sum_{j,k_1,k_2,m} \alpha_{j,k_1,k_2,m} b_{j,k_1,k_2,m}(x_1, x_2), \tag{2}$$

where $0 \le j \le J$ (indexing of scale), $0 \le k_1, k_2 < 2^j$ (indexing of localization), $0 \le m \le M(S(k_1, k_2, j))$ (indexing of direction) and $j, k_1, k_2, m \in \mathbb{N}$, $\alpha_{j,k_1,k_2,m} \in \{0,1\}$, $b_{j,k_1,k_2,m} \in B$.

The beamlet indexing of direction (parameter m in the above definition) may be introduced in many ways. The one often applied uses the polar system of coordinates to parameterise the beamlet direction and localization within the dyadic square.

From the practical point of view the transform works in the following way. Assume that we have all edges extracted from the image. To such an image we use the beamlet transform so that having any dyadic square (we start this process with the whole image as the initial square) we must decide whether the edge lying within this square may be approximated (with zero error) by any beamlet from the beamlets dictionary or not. If the edge may be approximated it means that the edge is simple straight line within this square. Otherwise the quadtree of this square is performed and the process is repeated with all four subsquares separately. Finally, after finite number of steps, we achieve the whole perfect decomposition of the image. The quadtree partition may be performed till the considering square will be of one pixel size. Further division is not possible.

In [17] the following definition of feature extractor called the i2D Selective Beamlet Operator (i2D SBO) based on beamlets has been proposed

$$F_B(x_1, x_2) = \sum_{J,k_1,k_2,m} \alpha_{J,k_1,k_2,m} b_{J,k_1,k_2,m}(x_1, x_2), \tag{3}$$

where $J = \lg N$, $0 \le k_1, k_2 < 2^j$, $0 \le m \le M(S(k_1, k_2, j))$ and $\alpha_{J,k_1,k_2,m} \in \{0,1\}$, $b_{j,k_1,k_2,m} \in B$. The operator is based on the fact that from the intrinsic dimensionality point of view all straight fragments of edges (the i1D ones) are approximated with 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

In our experiments many images of human faces have been investigated. The face's images are freely available from [22]. We have used the method of i2D feature extraction based on the geometrical wavelets approach, described in the previous section, due to the wide possibility of some parameters changing, which allow for adjusting the extractor to obtain the best results. Because all the methods described above give comparable results we limited our experiments to the only one method. The comparison of working of many feature extractors is beyond the scope of our paper.

In Fig. 2 a few sample tested images of human faces are presented. In Fig. 3 extracted edges of the original images are shown. To avoid much false edge detection the images have been smoothed first and then the detected edges have been binarised. Finally, in Fig. 4 i2D features of presented images of faces are shown, of course also in binary form.

One can see that removing from images i0D information, that is textures (see Fig. 3), the faces are still quite well visible. But note that if only i2D information is present (Fig. 4) the faces are also quite well visible. Moreover, they are distinguishable.

Fig. 2. Original images of faces (i0D+i1D+i2D signals).
Reproduced with permission from AT&T Laboratories Cambridge [22].

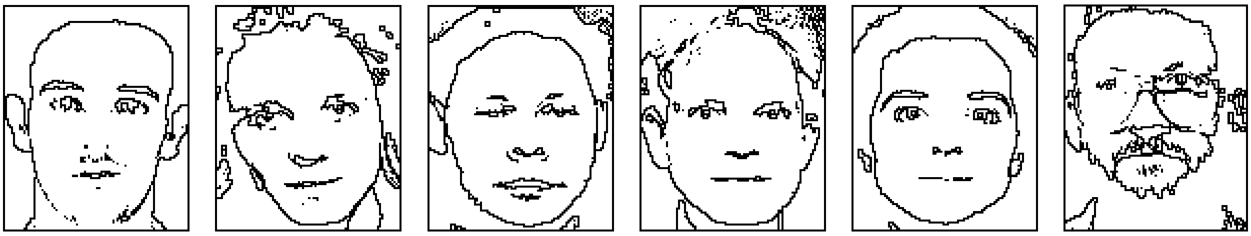

Fig. 3. Edges of faces (i1D+i2D signals).

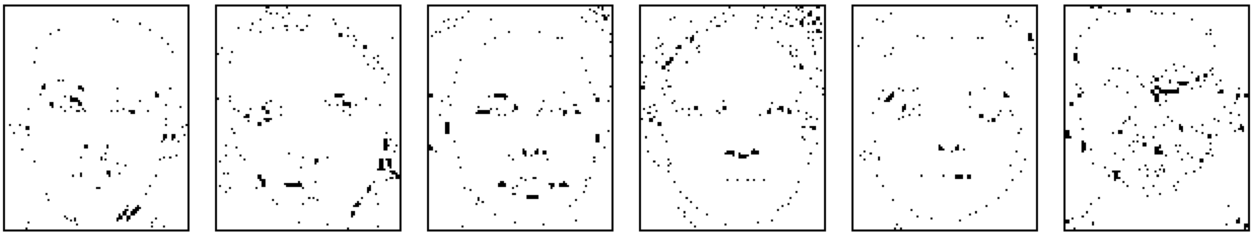

Fig. 4. i2D features of faces (i2D signal).

For further comparison purposes in Tab. 1 the numbers of pixels that form Fig. 3 and Fig. 4, respectively are presented. Additionally, in the last row of the table the ratio of the number of edge pixels per number of i2D pixels of each image are given. As one can see, the i2D information makes up only about 20% of information of all edges. It is astonishing that basing only on so few pixels one can build an image, on which not only human face is visible but also one may distinguish different faces. For comparison purposes note that the original images of faces are of size 92 × 112 pixels what gives 10304 pixels in total.

Table 1. Number of pixels of consecutive images from Figs. 3, 4.

Img	1	2	3	4	5	6
Edges	856	1229	1080	1204	869	1367
i2D	176	254	231	276	141	322
Ratio (%)	20,6	20,7	21,4	22,9	16,2	23,6

5. CONCLUSIONS

In the paper the influence of psychology of vision on the way, in which a man percepts images from surrounding world was presented. Because not all information is obtained at first look, only the more important features from image go to human brain. The results of carried out experiments are promising. They proved that the extractor based on geometric wavelets works very well and very good simulates behaviour of human brain during image reconstruction. It is unbelievable that from a very small amount of information the reconstruction of human face for identification purposes 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 fact is very important from database point of view.

It is also worth to mention that theory of geometrical wavelets is not complicated from mathematical point of view. It means that the extractor based on geometric wavelets is characterized by comparatively high speed in extraction features from images because of small complexity of the algorithm.

The authors think that it is possible to work out an automatic method for human face recognition and identification basing on the approach presented in the paper. But it requires further investigations. Further directions of investigations that are closely related to medicine are possible to establish. For example, basing on the proposed feature extractor one can try to use it for determining of the emotional states of a person basing on face images. Really, it is well known that if somebody is happy or is sad then his emotional state can be recognised from his/her face. Similarly, there is a high correlation between the level of pain occurring during many diseases and the appearance of the patient's face. Therefore it should be also possible to apply the feature extractor in those situations.

BIBLIOGRAPHY

[1] ATTNEAVE F., Some Informational Aspects of Visual Perception, Psychological Review, Vol. 61, pp. 183-193, 1954.

[2] BARTH E., ZETZSCHE C., KRIEGER G., Curvature Measures in Visual Information Processing, Open Systems and Information Dynamics, Vol. 5, pp. 25-39, 1998.

[3] BEAUDET P. R., Rotational Invariant Image Operators, International Joint Conference on Pattern Recognition, pp. 579-583, 1978.

[4] BIEDERMAN I., Recognition-by-Components: A Theory of Human Image Understanding, Psychological Review, Vol. 94, No. 2, pp. 115-147, 1987.

[5] BIEDERMAN I., COOPER F. E., Priming Contour-Deleted Images: Evidence for Intermediate Representations in Visual Object Recognition, Cognitive Psychology, Vol. 23, pp. 393-419, 1991.

[6] BIGÜN J., GRANLUND G. H., Optimal Orientation Detection of Linear Symetry, IEEE First International Conference on Computer Vision, Great Britain, pp. 433-438, 1987.

[7] BLAKEMORE C., OVER R., Curvature Detectors in Human Vision?, Perception, Vol. 3, 1974.

[8] CANDÈS E., DONOHO D., Curvelets – A Surprisingly Effective Nonadaptive Representation for Objects with Edges, Curves and Surfaces Fitting, A. Cohen, C. Rabut, and L. L. Schumaker, Eds. Saint-Malo, Vanderbilt University Press, 1999.

[9] DONOHO D. L., Wedgelets: Nearly-Minimax Estimation of Edges, Annals of Statistics, Vol. 27, pp. 859-897, 1999.

[10] FÖRSTNER W., GÜLCH E., A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features, ISPRS Intercommission Workshop, Interlaken, pp. 149-155, 1987.

[11] GALLANT J. L., BRAUN J., ESSEN D. C. V., Selectivity for Polar, Hyperbolic and Cartesian Gratings in Macaque Visual Cortex, Science, Vol. 259, pp. 100-103, 1993.

[12] 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.

[13] KAISER P. K., The Joy of Visual Perception, http://www.yorku.ca/eye/thejoy.htm .

[14] KOBATAKE E., TANAKA K., Neuronal Selectivity to Complex Object-Features in the Ventral Visual Pathway of the Macaque Cerebral Cortex, Journal of Neurophysiology, Vol. 71, pp. 856-867, 1994.

[15] 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.

[16] KRÜGER N., FELSBERG M., A Continuous Formulation of Intrinsic Dimension, Proceedings of the British Machine Vision Conference, 2003.

[17] LISOWSKA A., Intrinsic Dimensional Selective Operator Based on Geometrical Wavelets, Journal of Applied Computer Science, in review, 2004.

[18] MAKIEŁA M., Methods of faces' image identification based on characteristic feature points, MSc Thesis, Institute of Informatics, University of Silesia, Sosnowiec, (in Polish), 2002.
[19] 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.
[20] MITRA S. K., SICURANZA G. L., Nonlinear Image Processing, Academic Press, San Diego, 2001.
[21] SMITH S. M., BRADY J. M., Susan – a New Approach to Low Level Image Processing, International Journal of Computer Vision, Vol. 23, No. 1, pp. 45-78, 1997.
[22] THE DATABASE OF FACES, http://www.uk.research.att.com/facedatabase.html .
[23] 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.