

# Pseudo-entropy Similarity for Human Biometrics

Leonid Kompanets, Janusz Bobulski, Roman Wyrzykowski

Technical University of Czestochowa  
Institute of Mathematics and Computer Science  
Dąbrowskiego str., 73, 42-200, Czestochowa, Poland  
leonidfm@matinf.pcz.czyst.pl

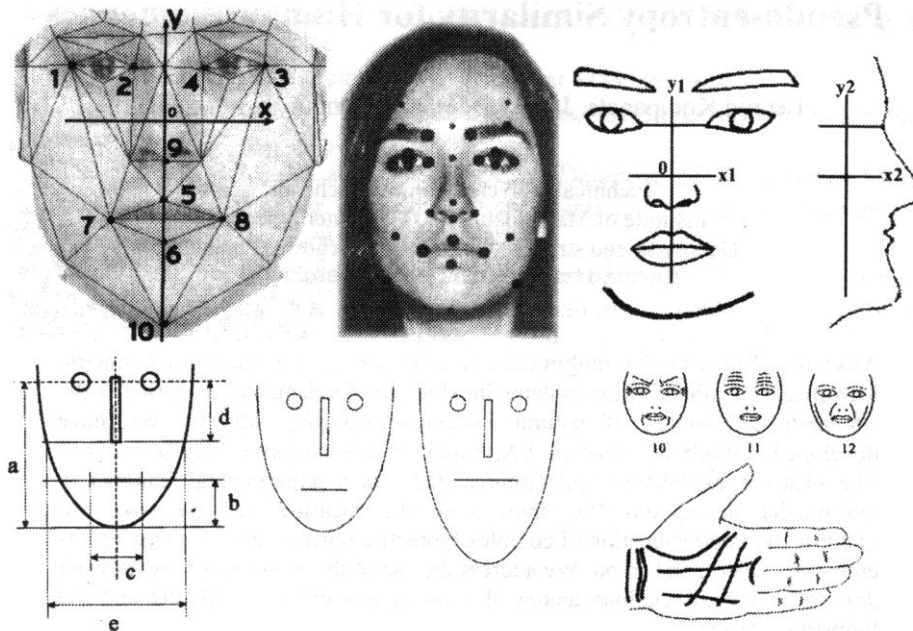
**Abstract.** With complex multimedia data, we see the emergence of biometric identification/authentication systems in which the fundamental operation is the *similarity assessment* of natural information-carrying objects. We have developed a similarity measure *JeK*, based on new notion of *pseudo-entropy*. The measure exhibits several features that match experimental findings in multimedia perception. We show how the measure can be used for identification/authentication of complex biometric objects, such as faces and its emotions, voices, and so on. We address the use of the pseudo-entropy measure *JeK* to deal with relations among the varied properties of 1D, 2D and 3D biometric objects.

## 1. Multimedia and Artificial Intelligence Background

Modern state of development of multimedia and artificial intelligence technologies, which also include biometric technologies, triggers off some new theoretical problems, which can not have effective solutions by means of the classic mathematics and informatics knowledge. Measuring of meaningful similarity of two natural biometric objects for the situations shown in Fig.1-6 is an example of such problems. Obviously, on the one hand, as shown in Fig.1 and 2 face approximations by means of the facial locations [6] or the map of frequency saliencies [4], are rough enough. On the other hand, the traditional measures of nearness, for example, shown in [1, p.419; 3, 8], are not adequate to any psycho-physiology measures of natural multimedia objects that are used by human beings.

In our opinion, the fundamental reason, why such situation exists, connects with the phenomenon that is described by Fechner's and Weber's laws. According to the laws, an intensity of a perceptual impression is proportional to a value logarithm of physical stimulus of a sense receptor, and a sensitivity threshold of any receptor is proportional an intensity of the stimulus. It means that similarity properties of any physical spaces (for example, colorimetric one) are completely different to the ones of psycho-physiology spaces of human being.

To measure meaningful similarity of two objects, it is necessary to find a set of features which adequately encodes the object characteristics that we intend to measure and endow the feature space with a suitable metric.



Figures 1 [1, background], 2 [7], 3 [8, background], 4 [9], 5, 6

To select the right set of features, characterize an objects as a points in a suitable vector space, researchers make some uncritical assumptions about the metric of the space. Typically, the feature space is assumed to be Euclidean. In mathematical statistics and pattern recognition theory, it was worked some theoretical,  $n$ -dimensional, Gaussian models that possess the separability property. It means that the model, by means of a hiperplane construction, can be separate on simple (or marginal) models. Such procedure is an instrument to construct some classify or recognize object procedures in terms of the 1 and 2 error types. The simplest model of the Gaussian mixture is Fisher's model.

Analytical measure to separate  $n$ -dimensional models is Machalanobis's measure. Euclidean, weighted Euclidean and Hamming distances are particular case of Machalanobis's type of metric. Generalization of Machalanobis's distance on arbitrary probability density function (*pdf*) is Kullback-Leibler's divergence [8, 9, 10] or symmetrized G. Jeffrey's divergence  $J^{Sh}$  between *pdf*  $p_x$  and  $p_y$  of two stochastic values  $X$  and  $Y$  with a generalized area existence  $Z$ :

$$J^{Sh} [X, Y] = \int_Z dz [p_x(z) - p_y(z)] \log_2 \frac{p_x(z)}{p_y(z)} \quad (1)$$

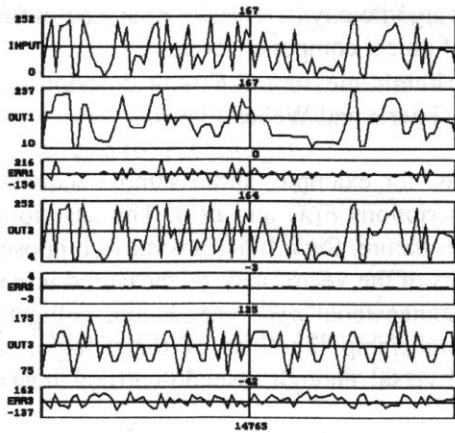


Figure 7

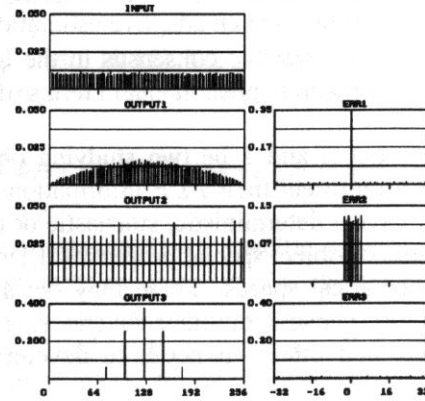


Figure 8

## 2. Statement of Problem

Advantages of G. Jeffrey's divergence by comparison with Shannon's,  $\chi^2$ -, Heffding's, Hellinger's, Renyi's, Rathie's types of the divergences (also entropy, information) [2, 8-12] and traditional measures [1, 3] are its logarithmic essence as well as the rich and understandable to an engineer the content of the notion of Shannon's global entropy. Disadvantages of G. Jeffrey's divergence are only statistical interpretation, lacking of values normalization, analog notation, lacking of the simplest intellectual procedures of measuring [8-12].

In spite of the rare applicability, G. Jeffrey's divergence has been selected as basic measure to synthesis of the *JeK* measure similarity. The measure *JeK* proposed and partly studied in [9, 10]. Here, the measure has been oriented on the similarity measuring in the case of non-traditional complex multimedia objects.

As it has been shown below, the measure *JeK* possess the new useful features:

- A) Universality with reference to nature and dimension of functions. A pair of any nature comparing functions must be transformed to the *PDF*-functions (see (5, 6)).
- B). Possibility measuring of the vector of similarity for sample sequences, spectrums, ect.
- C). Normalization of similarity value  $J^{Sh}$  (in % or %%) by the entropy  $H(X)$  and  $H(Y)$  objects.
- D). Content-related interpretation of the similarity in terms of the statistical theory of information, that are understandable to an engineer community.
- E). Sensitivity to any structural change of the *PDF*-functions.
- F). Digital notation.
- G). Possibility of fitting of specific reference points of comparing objects on a monitor.

70 Leonid Kompanets, Janusz Bobulski, and Roman Wyrzykowski

- H). Possibility of similarity measuring into any (Cartesian, polar) co-ordinate systems and an adaptive estimation of object component weights.
- I). Possible consensus in the logarithmic measure  $JeK$  with descriptions of perceptual phenomena in terms of Fechner's and Weber's laws.

Let  $X$  and  $Y$  be two studying objects, for example, audio sample sequences, which represent their  $PDF$ -information descriptions  $F(X)$  and  $L(Y)$ . The descriptions may have a deterministic, stochastic or other nature. Such functions may compose by vectors of object special properties. Elements of the vector may be measured into any mathematical spaces. These may be: a  $n$ -dimensional  $pdf$ -; correlation, colligation, spectral, cepstral or other functions, a membership functions of fuzzy logic; time series, and others. It needs to develop universal enough, pseudo-entropy measure ( $PEM$ )

$$PEM [F(X), L(Y)] = PEM [PDF(F_1, L_1), \dots, PDF(F_N, L_N)]. \quad (2)$$

which associates with an *integral information distance* between  $X, Y$ .

The measure also must satisfy the above-cited features A-I. *It may also be used to estimate a fusion (integration) effect in any multi-biometric system.*

### 3. Notions of Pseudo-entropy, Pseudo-information, Pseudo-divergence

The values of entropy, statistical information quantity and divergence are calculated on a base of the well-known functionals with the  $pdf$ -functions as their arguments. The  $pdf$ -functions possess some special peculiarities: a function values are located into  $[0,1]$  interval; an area under the function equals 1.

The functionals have a stochastic interpretation exceptionally. To *extend the fundamental contents of the notions of entropy, stochastic information, and divergence to the non-stochastic events*, it has used, as the arguments in (1), instead of the  $pdf$ -functions, another functions called the  $PDF$ -functions. These have also possessed the peculiarities like the  $pdf$ -functions. After such preprocessing of any nature function, the statistical theory relations preserve and guarantee the correctness of an extended interpretation of the notions of pseudo-entropy, statistical pseudo-information, pseudo-divergence like in the case of stochastic events.

Because of the generalization of the function nature, Shannon's entropy functional, calculated with the  $PDF$ -arguments, have been named the PSEUDO-ENTROPY. Then, the pseudo-divergence, calculated with the  $PDF$ -argument, may be interpreted as the absolute information distance between a pair of the objects. To receive a *relative information distance*, we must normalize the absolute distance by the sum of two objects entropy values. When the unit of the entropy is *Shannon* [Sh] (or [dit], [nit]), then the unit of the relative information distance is [%] or [%%]. *Concretization of the PDF-argument nature gives the possibilities to measure the a complex relative information similarity of an objects pair.*

#### 4. Mathematical and Algorithmic Notations of the *JeK* Criterion

It was created some measures (*KoD*, *KoT*, *KoJ*, *JeK*,) [8-11]. We discuss the measure *JeK* features. Generalized mathematical and algorithmic notations of the measure given (5) and (6) accordingly.

Into (5),  $\tilde{Hist}[\cdot]$  are the digital ( $\perp\perp$ ) *PDF*-functions (pseudo-histograms) of two information-carrying objects mapped on the grid *I*.

Mark „ $\pm$ ”, situated before the *JeK*[.] value, characterize a specific relationship between *F*- and *L*-function. Accordingly an assumption, function *F* is the reference one. If the values of *F* and *L* functions equal, then the mark is neglected.

A denominator in (5) and (6) includes a sum of Shannon’s global ( $q=1$ ) pseudo-entropies. It is a component for the normalization of G. Jeffrey’s divergence value which is situated in a numerator.

Component “ $\Sigma\{\cdot\}$ ” in (6) is an algorithm to calculate G.Jeffrey’s divergence. Component “ $\Sigma[\cdot]$ ” is an algorithm to calculate Shannon’s global entropies  $H[F_i(i)]$ ,  $H[L_{i\pm m}(i)]$ .

The *JeK* measure is calculated for a given grid *I*, which has a number *I* of the *i*-th elements.  $F_i(i)$ ,  $L_{i\pm m}(i)$  functions are transformed to form of the *PDF*-functions as:

$$\sum_i f_i(i) = 1, \quad \sum_{i\pm m} l_{i\pm m}(i) \leq 1; \quad f(i) \text{ and } l_{i\pm m}(i) \in [0,1] \quad (3)$$

Into the interval  $[0, >1]$ , a degree of the measure *JeK* sensitivity depends on the ratio:

$$I \times \sum_i |F_i(i) - L_{i\pm m}(i)| \quad (4)$$

Index ‘*i*’ is an element of a grid *I*. Index ‘ $\pm m$ ’ characterizes a direction ( $\pm$ ) and a value (*m*) of function *L* displacement relative to function *F*.

In the event of 2D and 3D objects, the 2G property may be achieved by means of designing a ( $\pm m, n, p$ )-index set for (4) and an automatic procedure of meaningful indexes manipulation.

$$\pm JeK_{\perp\perp}^{Sh} [\tilde{Hist}(\hat{Ob}_1, \check{Ob}_2)] = \frac{\sum_{i=1}^I (\hat{P}_{1i} - \check{P}_{2i}) \cdot \ln(\hat{P}_{1i} / \check{P}_{2i})}{-\sum_{i=1}^I \hat{P}_{1i} \cdot \ln \hat{P}_{1i} - \sum_{i=1}^I \check{P}_{2i} \cdot \ln \check{P}_{2i}} \quad (5)$$

$$\pm JeK_{\perp\perp}[F_i(i), L_{i\pm m}(i)] = \frac{\sum_{i,j\pm m(i=1,I,m=-I,I)} \left\{ \begin{array}{l} [f_i(i) - l_{i\pm m}(i)] \ln \left[ \frac{f_i(i)}{l_{i\pm m}(i)} \right], \\ \text{if } 0 < [f_i(i) \text{ and } l_{i\pm m}(i)] < 1 \\ i = 1, \dots, I; m = -I, \dots, 0, \dots, I; \\ - [f_i(i)] \ln [f_i(i)], \\ \text{if } 0 < [f_i(i)] < 1 \text{ and } [l_{i\pm m}(i)] = 0 \text{ or } 1, \\ i = 1, \dots, I; m = -I, \dots, 0, \dots, I; \\ - [l_{i\pm m}(i)] \ln [l_{i\pm m}(i)], \\ \text{if } 0 < [l_{i\pm m}(i)] < 1 \text{ and } [f_i(i)] = 0 \text{ or } 1, \\ i = 1, \dots, I; m = -I, \dots, 0, \dots, I; \\ 0 \\ \text{if } [l_{i\pm m}(i)] \text{ and } [f_i(i)] = 0 \text{ or } 1, \\ i = 1, \dots, I; m = -I, \dots, 0, \dots, I; \\ 0, \\ \text{if } [l_{i\pm m}(i)] = [f_i(i)], \\ i = 1, \dots, I; m = -I, \dots, 0, \dots, I; \end{array} \right.}{-\sum_{i(i=1,I)} [f_i(i)] \ln [f_i(i)] - \sum_{i,j\pm m(i=1,I,m=-I,I)} [l_{i\pm m}(i)] \ln [l_{i\pm m}(i)]} \quad (6)$$

The measure depends on a scale of the grid  $I$  and it is sensitive to changing of a signal structure (for example, to changing a number of gray-levels of the functions  $F_i(i)$ ,  $L_{i\pm m}(i)$ ).

The measured value changes in the  $[0, >1]$ -interval. The value  $=0$  if  $F_i(i) \equiv L_{i\pm m}(i)$  and  $m=0$ . The value  $\geq 1$  if a function form transforms to the form of the  $\delta$  (delta) function or if it executes the condition  $\pm m \geq \pm I$  for  $L_{i\pm m}(i)$  function.

The  $JeK$  measure of two objects similarity has designed on the base of formula (3). The measure fulfils the conditions  $A - I$ .

## 5. Example of Automatic Measuring of Audio Objects Similarity

We present some results on an experimental estimation of the measure  $JeK$  sensitivity to a signal *structure variation* (the feature  $2E$ ).

To do it, white noise sample sequence (Input) was processed by median filter (Out1, Err1), Walsh's filter of constant component (Out2, Err2), and quasi-inverted compressor (Out3, Err3) with compression ratio  $Ksq$ . A posterior values of standard deviations  $\sigma$ , normalized pseudo-divergence  $JeK$  values, fragments of sample sequences and their histograms, which include information about signal structures, shown in Table 1, Fig.7, 8.

**Table 1**

Signal type	Transformation characteristics	Median filtration (aperture 3)	Walsh filtration (aperture 4)	Quasi-inverted compression
White noise	$Ksq [X]$	1.00	2.67	64.00
- "-	$\sigma[\%]$	<b>26.25</b>	<b>0.93</b>	<b>26.70</b>
- "-	$JeK [\%]$	<b>3.60</b>	<b>80.9</b>	<b>102.4</b>

## 6. Sensitivity, Stability and Relative Linearity of Measure $JeK$ in Area of Similarity

To get acquainted with the above-cited measure  $JeK$  properties (especially with the  $2G$  one), let us get familiar with the results of the second experiment. Fragments of audio-signal sequences which have sample size 1024, 256, 512 accordingly, and their histograms shown in Fig.9-13. The samples are encoded by 8 bits binary code.

In the window of Fig.9, it showed a specific result of measure  $JeK[.]$  calculating in according to formula (6) for the case of two equal and not shifted signals which have sample size 1024. In Fig.10-11 and 12-13, there is shown an effect of shifting of the sequence  $L$ , for the sequence fragments which have sample size 256 and 512 accordingly. The behavior of the measure  $JeK[.]$  represented for cases of variation of difference values  $D(F, L)$  between functions  $F, L$  (Fig.14, Table 2) and of variation of displacement  $\pm m$  function  $L$  (Fig.15, Table 2).

For the case of near-situated pair of the functions, experiment results confirm high sensitivity, stability and linear enough behavior of the  $JeK$  measure. In the case of a given grid  $I$ , including 256 and 512  $i$ -th indexes, a linear area of the  $JeK$  measure equals  $0,0 - 0,8$ ; and a figure superimposing procedure has linear type into an area of  $\pm 40$  samples.

## 7. Examples and Perspective of Similarity Measuring of 2D and 3D Biometric Objects

It imagines that the new  $(A - I)$  features of the  $JeK[.]$  measure give effective possibility of measuring the similarity in the case of 2D and  $nD$  biometrics. The examples of typical objects shown in Fig.1-6. Peculiarities of the objects are their non-traditional complex ("non-mathematical") form and little enough differences of the forms. Also, it creates possibility of premeasuring procedures design to combine the centers of co-ordinate systems of objects pair on monitor.

In Fig.1, 3, an example of the figure center finding shown, for the event of face expression similarity measuring. When measuring, it may be accepted, for example, that a distance between eye corners equals 60 pixels. The center of co-ordinate system

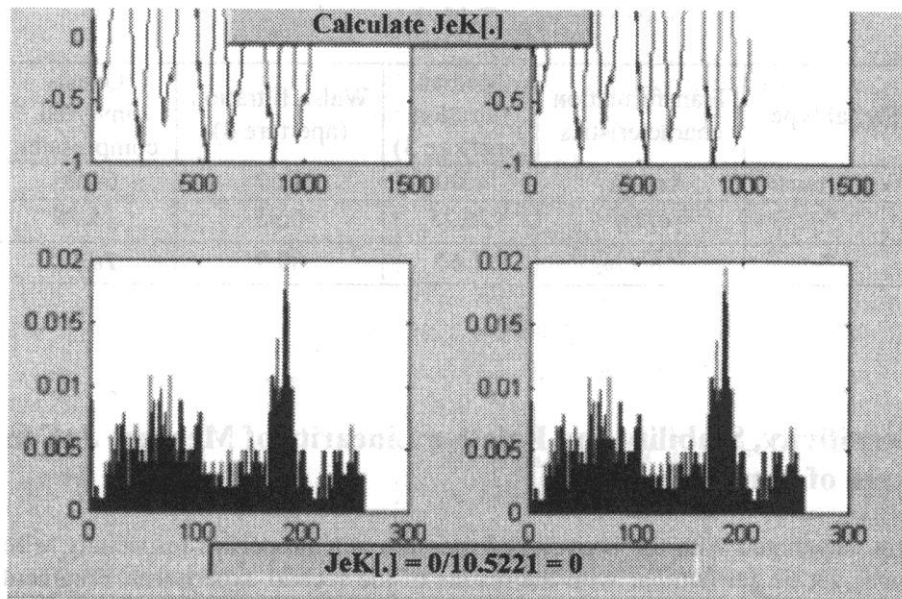


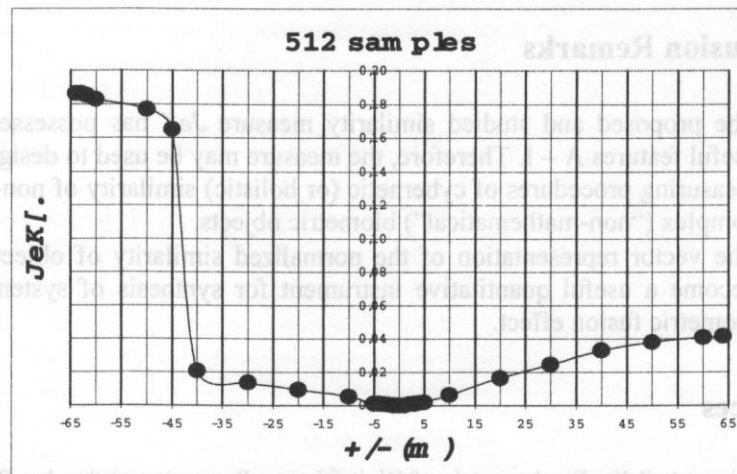
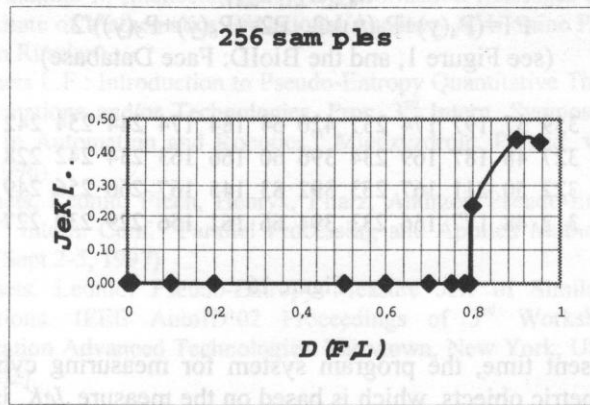
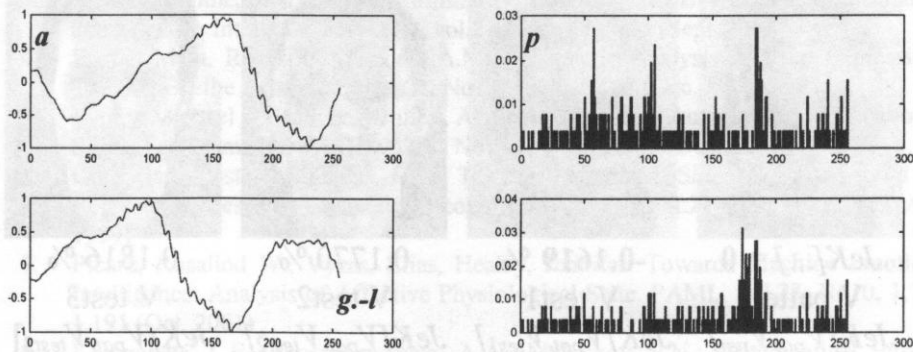
Figure 9

Table 2

256 samples		512 samples			
$D(F,L)$	JeK %	$-m$	JeK %	$+m$	JeK [.]%
0.006	0.0000	-1	0.0090	0	0.0000
0.007	0.0441	-2	0.0301	1	0.0107
0.1	0.0441	-3	0.0544	2	0.0584
0.2	0.2392	-5	0.1107	3	0.1006
0.5	0.2392	-10	0.5074	5	0.1614
0.6	0.0441	-20	9.3276	10	0.6411
0.7	0.2392	-30	13.398	20	1.6322
0.79	0.0253	-45	16.478	30	2.4267
0.8	23.221	-50	17.665	50	3.8082
0.95	43.305	-64	18.595	64	4.1396

may also be situated on 60 pixels below as shown in Fig.1 3. In this case, instead of a Cartesian co-ordinate system may be utilized a polar system in which are manipulated by  $2^x$  angles (as in the case of FFT). Notice that the preprocessing procedures of biometric information are state-of-the-art.





Figures 10, 11, 12, 13, 14, 15



$$\begin{array}{cccc}
 JeK[.,.] = 0 & -0.1619 \% & 0.1770 \% & 0.1816 \% \\
 V_{pattern} & V_{test1} & V_{test2} & V_{test3} \\
 JeK[V_{pat}, V_{pat}] & JeK[V_{pat}, V_{tes1}] & JeK[V_{pat}, V_{test2}] & JeK[V_{pat}, V_{test3}]
 \end{array}$$

$$\text{Vector } V[1...16] = [d_{2,4}; d_{1,3}; d_{5,6}; d_{7,8}; d_{9,P1}; d_{P1,P2}; d_{10,P1}; d_{9,P2}; d_{10,P2}; d_{9,P1}; d_{7,1}; d_{8,3}; d_{7,2}; d_{8,4}; d_{2,1}; d_{3,4}].$$

$$P1 = (P_4(y) + P_2(y)) / 2. \quad P2 = (P_6(y) + P_5(y)) / 2.$$

(see Figure 1, and the BioID: Face Database)

$$\begin{array}{l}
 V_{pattern} = [120 \ 359 \ 37 \ 197 \ 174 \ 237 \ 420 \ 64 \ 184 \ 174 \ 244 \ 254 \ 242 \ 252 \ 125 \ 115]. \\
 V_{test1} = [118 \ 327 \ 48 \ 187 \ 169 \ 234 \ 396 \ 60 \ 166 \ 163 \ 234 \ 242 \ 228 \ 250 \ 114 \ 105]. \\
 V_{test2} = [123 \ 322 \ 50 \ 211 \ 167 \ 253 \ 392 \ 83 \ 145 \ 167 \ 260 \ 256 \ 249 \ 260 \ 112 \ 97]. \\
 V_{test3} = [106 \ 312 \ 48 \ 177 \ 156 \ 233 \ 394 \ 66 \ 161 \ 156 \ 229 \ 239 \ 227 \ 241 \ 103 \ 104].
 \end{array}$$

Figure 16

In the present time, the program system for measuring cybernetic (holistic) similarity of biometric objects, which is based on the measure  $JeK$ , is designed by the author.

## 8. Conclusion Remarks

- The proposed and studied similarity measure  $JeK$  has possessed the new useful features A – I. Therefore, the measure may be used to design effective measuring procedures of cybernetic (or holistic) similarity of non-traditional complex (“non-mathematical”) biometric objects.
- The vector representation of the normalized similarity of object pair may become a useful quantitative instrument for synthesis of system used the biometric fusion effect.

## References

1. Jain, Anil K.: Fundamentals of Digital Image Processing. 4-th edn., Prentice-Hall International, Inc.
2. Vajda, J.: On the  $f$ -Divergence and Similarity of Probability Measures. Period. March. Hung., v.2, 223-226 (1972)

3. Santini, Simone, Jain, Ramesh: Similarity Measures. IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), vol.21, No9, 871-883 (Sept 1999)
4. Pantic, Maja, Rothkrantz, Leon J.A.M.: Automatic Analysis of Face Expressions: The State of the Art. PAMI, vol.22, No12, 1,425-1,445 (Dec. 2000)
5. Lyons, Michael J., Budynek, Julien, Akamatsu, Shigeru: Automatic Classification of Single Facial Images. PAMI, vol.21, No12, 1,1357-1,362 (Dec. 1999)
6. Craw, Ian, Costen, Nicholas, Kato, Takashi, Akamatsu, Shigeru: How Should We Represent Faces for Automatic Recognition?. PAMI, vol.21, No8, 725-736, (Aug. 1999)
7. Picard, Rosalind W., Vyzas, Elias, Healey, Jennifer: Towards Machine Emotional Intelligence: Analysis of Affective Physiological State. PAMI, vol.23, No10, 1,175-1,191 (Oct. 2001)
8. Kompanets, L.F., Krasnoproslyna, A.A., Malyukov N.N.: Mathematical Base of Research in Automatics and Control. Vyscha shkola, Kiev (1992, In Russian)
9. Kompanets L. F.: The computer Criterion *JeK* for Quality Estimating of Information Transformation. In Intellectualization of Information Messages Processing Systems. The Institute of Mathematics of National Academy of Ukraine Press, No 1, 101-106 (1995, In Russian)
10. Kompanets L.F.: Introduction to Pseudo-Entropy Quantitative Theory of Information Transformations and/or Technologies. Proc. 3<sup>rd</sup> Intern. Symposium. "Methods and Models in Automation and Robotics," Miedzzydroje, Poland, vol.2,665-670 (Sept. 10-13, 1996)
11. Kompanets, Leonid, Piech, Henryk, Pilarz, Andrzej: Pseudo-Entropy and Beyond. Proc. 2<sup>nd</sup> Intern. Conf. "Parallel Processing and Applied Mathematics," Zakopane, Poland (Sept 2-5, 1997)
12. Kompanets, Leonid: Pseudo-Entropy Measure *JeK* of Similarity for Biometric Applications. IEEE AutoID'02 Proceedings of 3<sup>rd</sup> Workshop on Automatic Identification Advanced Technologies. Tarrytown, New York, USA, 142-146 (14-15 Mar. 2002).