2DLDA-based texture recognition in the aspect of objective image quality assessment

Krzysztof Okarma[1*], Paweł Forczmański[2*]

[1]*Chair of Signal Processing and Multimedia Engineering, Szczecin University of Technology, 26. Kwietnia 10, 71-126 Szczecin, Poland*
[2] *Institute of Computer Graphics and Multimedia Systems, Szczecin University of Technology, Żołnierska 49, 71-210 Szczecin, Poland*

Abstract

The image quality is a crucial property of each image when it comes to successful recognition. There are many methods of image quality assessment which use both objective and subjective measures. The most desirable situation is when we can evaluate the quality of an image prior to recognition.

It is well known that most of classical objective image quality assessment methods, mainly based on the Mean Square Error, are poorly correlated with the way humans perceive the quality of digital images. Recently some new methods of full-reference image quality assessment have been proposed based on Singular Value Decomposition and Structural Similarity, especially useful for development of new image processing methods e.g. filtration or lossy compression.

Despite the fact that full-reference metrics require the knowledge of original image to compute them their application in image recognition systems can be also useful. In the remote controlled systems where lossy compressed images are transferred using low bandwidth networks, the additional information related to the quality of transmitted image can be helpful for the estimation of recognition accuracy or even the choice of recognition method.

The paper presents a problem of recognizing visual textures using two-dimensional Linear Discriminant Analysis. The image features are taken from the FFT spectrum of gray-scale image and then rendered into a feature matrix using LDA. The final part of recognition is performed using distance calculation from the centers of classes. The experiments employ standard benchmark database – Brodatz Textures.

Performed investigations are focused on the influence of image quality on the recognition performance and the correlation between image quality metrics and the recognition accuracy.

1. Introduction

Development of objective image quality assessment methods in recent years have caused some new possibilities of their usage in a wide area of applications.

---

*Corresponding authors: *e-mail addresses*; okarma@ps.pl; pforczmanski@wi.ps.pl

The traditionally used full-reference image quality metrics, such as Mean Square Error, Peak Signal to Noise Ratio etc., require the usage of original image for comparison with the distorted one being assessed. For many applications it is an acceptable approach, in some cases the additional information about the quality of the image can be even included in the image file for utilization if the original image is unknown for the end-user (e.g. after lossy compression or transmission).

Nevertheless, the main problem has been the poor correlation of traditional metrics with the subjective evaluation performed by human observers. Regardless of some experiments with rather complicated vector metrics based on the models of Human Visual System, some new efficient methods have been proposed in the last few years. Despite their full-reference character the obtained results are much better correlated with the way humans perceive images.

Including the reliable information about the image quality in the image files makes it possible to use them in many image recognition and classification systems where such metrics can be used in forecasting the recognition accuracy in the presence of some distortions caused e.g. by lossy compression or low bandwidth transmission. However, proper usage of image quality data requires the verification of some new metrics towards their usefulness and correlation with the results obtained during recognition. Because many kinds of distortions reduce the recognition accuracy in various ways, the image quality index obtained for such distorted images should be well correlated with the obtained recognition rate.
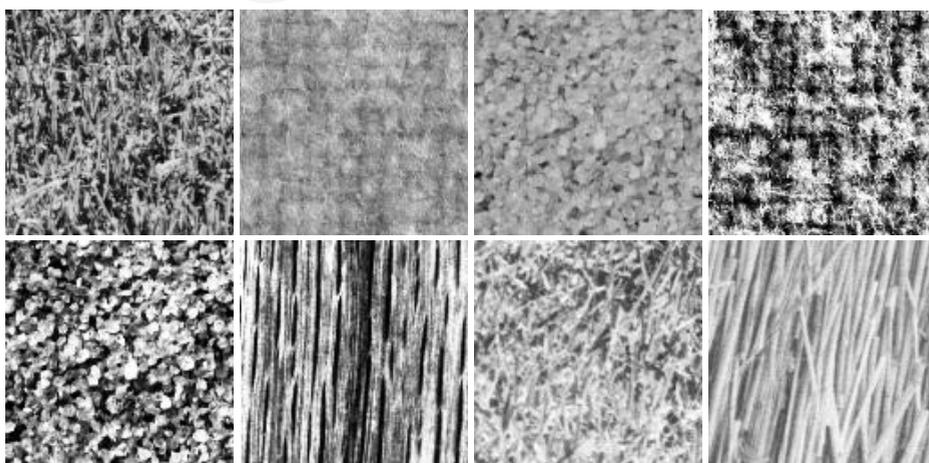

Fig. 1. Example images from Brodatz Album used in experiments

In this paper the usefulness of some modern image quality metrics to predict the visual textures' recognition accuracy in the presence of various distortions is

analyzed. As the textures' recognition method the two-dimensional Linear Discriminant Analysis algorithm is used. All the experiments are based on standard benchmark database of textures known as Brodatz Album [1]. Sample images used in the experiments are presented below in Fig. 1.

2. 2DLDA algorithm for texture recognition

The problem of automatic image recognition involves gathering, processing and storing large amounts of data. Reducing very high-dimensional feature spaces leads to speeding-up the computations. A method of Linear Discriminant Analysis (LDA) is commonly used for clustering of input data together with reduction of their dimensionality, hence providing much help when it comes to recognition. Classical LDA is applicable only when the number of images in the dataset is much larger than the dimensionality of a single image [2]. Moreover, using LDA for tasks such as image recognition or image retrieval can be challenging because it treats the data as one-dimensional, when in fact they are two-dimensional. That is why almost all presented algorithms involve some sort of dimensionality pre-reduction discarding in many cases the spatial relations between pixels.

One of the possible solutions of this problem is using two-dimensional transformation based on Principal Component Analysis (and Karhunen-Loeve Transform) and/or LDA. The first algorithm from this group was presented in [3], where a novel, two-dimensional version of KLT for face recognition task was developed. An extension of this method (as PCArc - for row and column representation) was presented in [4]. The same approach for LDA was used in [5], but the authors limited the application of LDA to the only one of image dimensions. Many previous publications show that two-dimensional LDA (2DLDA) can be applied for high-dimensional data [6,7]. It does not require any preliminary processing or additional reduction of dimensionality of input images. The experiments show also that the recognition rate is higher in comparison to other methods presented in literature. All these facts stay behind the application of 2DLDA for texture classification and recognition.

The algorithm of recognition discussed in this paper is composed of three basic elements. First, two-dimensional Fast Fourier Transform (2DFFT) is performed on each image giving a matrix containing spectral components selected from the lower part of the amplitude spectrum. Then it is reduced by means of 2DLDA producing final matrix of output features [7]. The last stage is the classification performed using the Euclidean distance calculation.

Let us assume that the whole dataset contains K classes of textures. Each class consists of L images. Each image $X^{(k,l)}$ contains $M \times N$ gray-scale pixels.

In the first step, for all images in the dataset we calculate global mean image $\overline{X}$ and mean image for each class $\overline{X}^{(k)}$. Then we remove the global mean image from each respective within-class mean:

$$\hat{X}^{(k)} = \overline{X}^{(k)} - \overline{X}, \ \forall k = 1,\ldots,K ; \tag{1}$$

and, a class-mean from each class member image, respectively:

$$\tilde{X}^{(k,l)} = X^{(k,l)} - \overline{X}^{(k)} \ \ \forall k = 1,\ldots,K, \ l = 1,\ldots,L . \tag{2}$$

In the next step we calculate two covariance matrices for both row representations of images:

$$W^{(R)} = \sum_{k=1}^{K}\sum_{l=1}^{L} \tilde{X}^{(k,l)}\left(\tilde{X}^{(k,l)}\right)^T , \tag{3}$$

$$B^{(R)} = \sum_{k=1}^{K} \hat{X}^{(k)}\left(\hat{X}^{(k)}\right)^T \tag{4}$$

and column representation, respectively:

$$W^{(C)} = \sum_{k=1}^{K}\sum_{l=1}^{L}\left(\tilde{X}^{(k,l)}\right)^T \tilde{X}^{(k,l)} , \tag{5}$$

$$B^{(C)} = \sum_{k=1}^{K}\left(\hat{X}^{(k)}\right)^T \hat{X}^{(k)} , \tag{6}$$

Then we build the total scatter matrices:

$$H^{(R)} = \left(W^{(R)}\right)^{-1} B^{(R)} , \tag{7}$$

$$H^{(C)} = \left(W^{(C)}\right)^{-1} B^{(C)} , \tag{8}$$

which correspond to the ratio of between-class scatter against within-class scatter. This ratio is known as the Fisher's criterion [2,8] and should be maximized.

For both $H^{(R)}$ and $H^{(C)}$ we calculate eigenvalue matrices $\{\Lambda^{(R)}, \Lambda^{(C)}\}$ (diagonal matrices of eigenvalues of $M \times M$ and $N \times N$ elements, respectively), and eigenvector matrices $\{V^{(R)}, V^{(C)}\}$ (orthogonal matrices of $M \times M$ and $N \times N$ elements, respectively, whose columns contain eigenvectors) [8]:

$$\Lambda^{(C)} = \left(V^{(C)}\right)^T H^{(C)} V^{(C)}, \tag{9}$$

$$\Lambda^{(R)} = \left(V^{(R)}\right)^T H^{(R)} V^{(R)}. \tag{10}$$

From the diagonals of $\Lambda^{(R)}$ and $\Lambda^{(C)}$ we select s largest elements and memorize their positions ($1 < s \le (K-1)$). From $\left(V^{(R)}\right)^T$ we select s rows corresponding to the s largest eigenvalues and from $V^{(C)}$ we select s columns, respectively. Then we build two new rectangular matrices: $F^{(R)}$ of $s \times M$ elements and $F^{(C)}$ of $N \times s$ elements, which will be used as the transformation matrices for 2DLDA. The transformation is performed according to the following equation [6,7]:

$$Y^{(k,l)} = F^{(R)}\left(X^{(k,l)} - \bar{X}\right)F^{(C)}, \tag{11}$$

giving a feature matrix *Y* for each block *X*.

The classification employs feature space created for all textures in the database. A query texture (image) is projected into it and the Euclidean distance is being calculated. Two alternative distances are taken into consideration: the first one is a distance to all elements, while the second one is the distance to the centers of all classes only. The latter approach is easier to perform and requires much less computations, however, its efficiency is lower in the case of irregular clusters. The selected recognition results of 3 experiments are presented in Fig. 2. They were performed in "FaReS-Modeller" Environment [9]. Each experiment gave 5 nearest images from the database for a query image (the first image in each row). The first row shows "perfect" recognition giving 5 correctly retrieved images. The second one shows correct answer on the first place only, while the remaining images are false. The last row shows the case when totally incorrect answer has been given. The distances are shown above each resulting image.
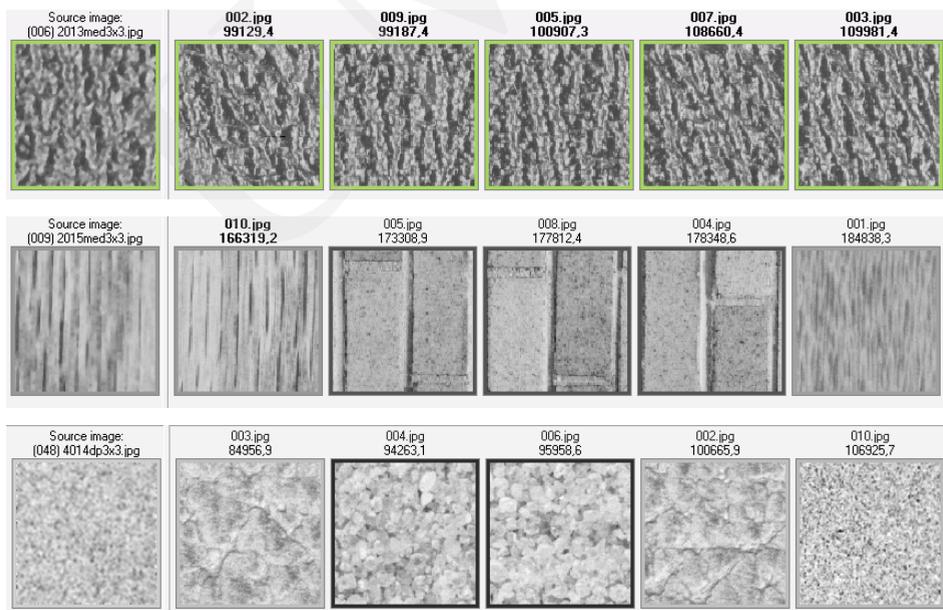

Fig. 2. Sample results of recognition (bold file names indicate proper recognition)

3. Modern techniques of objective image quality assessment

The idea of objective image quality assessment corresponds to the requirement of automatic description of the way humans perceive images processed using various algorithms, preferably using single scalar value, which

should be correlated as well as possible with the subjective evaluation by a typical observer. Such measures can be utilized in many areas of digital image processing e.g. development of new filtration algorithms, lossy compression methods or video transmission techniques.

The classical approach to objective image quality assessment is using Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR) and some other similar metrics briefly described in [10], which are unfortunately poorly correlated with Human Visual System (HVS). Some vector metrics based on HSV modelling, which have been also proposed, are less universal and too complicated for practical usage e.g. Picture Quality Scale (PQS) [11].

As the first successful attempt to the problem of universal image quality metric the proposition of Universal Image Quality Index [12] can be considered, further extended into Structural Similarity [13] defined as:

$$SSIM = \frac{\left(2 \cdot \overline{X} \cdot \overline{X'} + C_1\right) \cdot \left(2 \cdot \sigma_{XX'} + C_2\right)}{\left(\overline{X}^2 + \overline{X'}^2 + C_1\right) \cdot \left(\sigma_X^2 + \sigma_{X'}^2 + C_2\right)} , \qquad (12)$$

where: X and X' stand for the original and distorted images respectively, and the constants C_1 and C_2 are chosen such that they do not introduce any significant changes of the results (e.g. $C_1 = (0.01 \times 255)^2$ and $C_2 = (0.03 \times 255)^2$ as suggested by the authors of paper [13]. $\overline{X}$ and $\overline{X'}$ are the average values and σ denotes the standard deviation in the original and distorted image blocks respectively. The main role of C_1 and C_2 coefficients is prevention from dividing by zero especially for large flat and dark regions of the image.

Thanks to the SSIM index, it is possible to create quality map of the whole image since the formula is applied for image blocks using the sliding window approach. The default procedure proposed in [13] employs the Gaussian window for the blocks of 11×11 pixels. The average value of SSIM from the whole quality map is treated as the overall image quality metric, which is sensitive to three common types of distortions introduced by the image processing algorithms: loss of correlation, luminance distortion and loss of contrast.

Another idea, presented in [14] is using Singular Value Decomposition (SVD) applied for 8×8 pixels blocks of original and distorted images to compare them. For each block the following value is calculated:

$$D_k = \sqrt{\sum_{b=1}^{n} \left(s_b - s'_b\right)^2} , \qquad (13)$$

where s_b and s'_b are the singular values obtained for original and distorted image blocks ($n = 8$).

The overall quality metric is defined as the following expression:

$$M_{SVD} = \frac{1}{B} \cdot \sum_{i=0}^{B} |D_i - D_{mid}|, \qquad (14)$$

where B denotes the total number of blocks in the image and D_{mid} stands for the middle element of the sorted vector D.

Instead of using the SVD approach, the application of full image transforms (DFT, DWT and DCT) for quality assessment is also possible as presented in [15]. The idea of the methods used for all transforms is based on the calculations of the mean of four standard deviations of the differences between transform magnitude coefficients for four bands in the original and distorted images.

All methods discussed above belong to the group of full-reference ones and the necessity of using the original image limits their potential application area. However, that approach is the only one universal solution which leads to good correlation with subjective evaluation.

Subjective methods based on Mean Opinion Score require performing some tests where the observers evaluate the quality of presented images filling up the questionnaires and their statistical analysis is performed. They can be useful for the development of some new objective metrics, even better correlated with HVS, but cannot be directly used in computer applications where results should be calculated in the short time without any human interactions.

On the other hand, existing "blind" (no-reference) image quality assessment methods are still far from universality and perfection and need to be improved. The only worth noticing application is the JPEG artifacts measurement that can be successfully recognized without the usage of uncompressed original image [16].

For image recognition and classification purposes, there is often the situation of unknown "original" image and the only available information is the distorted image from the camera. In that case good "blind" methods would have a great advantage, but in the distributed systems, e.g. inspection ones, where data from many cameras should be lossily compressed and transmitted by relatively low bandwidth networks, using additional data related to the quality of compressed image (e.g. in the file header) can be very helpful. For highly compressed images or in the presence of some distortions, e.g. dependent on the lighting conditions, the prediction of possible recognition accuracy based on the image quality assessment seems to be promising. In some systems the additional estimation of the "differential" image quality metric for two neighboring video frames is also possible e.g. for detection of rapid changes of lighting conditions, image sharpness etc. and their influence on the recognition accuracy.

4. Influence of image distortions on the recognition accuracy and image quality

Each image recognition method, regardless of its purpose, is more or less sensitive on the quality of the input images. For most methods the dependency between quality of the images used in recognition process and the recognition accuracy can be described only in general: poor quality leads to weak recognition and better image quality should provide better results of recognition.

On the other hand, there are many image quality metrics which can be used for description of the image quality, not necessarily well correlated with the subjective assessment, but there is no guarantee that the recognition accuracy is correlated with such evaluation as well.

The idea of performed experiments is the evaluation of some recently proposed image quality metrics in order to check if they are correlated with the texture recognition accuracy using the 2DLDA algorithm described above.

Recognized textures have been distorted using several typical methods such as contamination by the impulse (salt and pepper) noise, lossy JPEG compression, low-pass filtration and median filtration. The example texture with various distortions is presented in Fig. 3. For each of 56 textures form the Brodatz database 21 images have been created as the different (sometimes overlapping) fragments of the texture, 10 of them have been used for training and remaining 11 images have been used for tests, also in their distorted versions. The average recognition accuracy for the images without any distortions is 92.21% (considering the distance to the nearest element) and 80.03% if the distance to the centers of classes has been used. These results are superior to those presented in literature [17]. The average results of recognition for different distortions as well as respective image quality metrics (also averaged for all textures used in the tests) are presented in Table 1.

Table 2 illustrates the correlation between the image quality metrics and the recognition accuracy for the "nearest element" and "centers of classes" approaches. Graphical illustration of the dependence between the most correlated metric and the recognition accuracy is presented in Fig. 4.

Analyzing the presented results we can notice that the correlation between each metric and the recognition accuracy is relatively low for the JPEG compressed images. All image quality measures considered in the paper are sensitive to the lossy JPEG compression while the recognition algorithm works fine even for strongly compressed images being almost insensitive to the distortions introduced in the JPEG compression. For additional verification of the results, correlation coefficients excluding lossily compressed images have been also calculated and the obtained results are presented in Table 3.
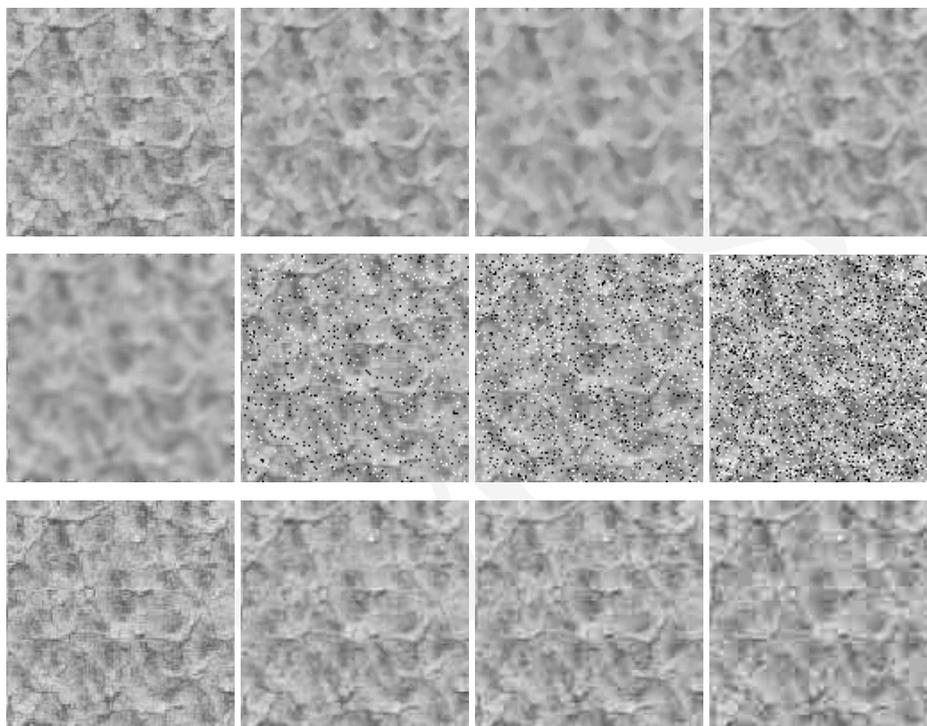
Fig. 3. Example original image and its distorted variants: median filtered (mask 3×3 and 5×5), low-pass filtered (mask 3×3 and 5×5), contaminated by the impulse noise (5%, 10% and 20% of pixels) and JPEG compressed (quality: 60%, 40%, 20% and 10%) respectively

Table 1. Average recognition accuracy and the averaged quality measures for textures with various distortions

Distortion	Recognition accuracy		Image quality					
	Nearest element	Centers of classes	SSIM	M_{SVD}	M_{DCT}	M_{DFT}	M_{DWT}	PSNR
Median 3x3	81.33%	71.59 %	0.8359	13.042	8.101	954.7	8.828	27.445
Median 5x5	62.18%	57.79 %	0.6614	22.570	11.407	1578.0	12.205	24.226
Low-pass 3x3	71.47%	63.37 %	0.7638	51.455	18.719	2912.5	13.658	19.150
Low-pass 5x5	46.83%	47.32 %	0.5302	50.476	31.319	5224.3	21.507	14.131
5% impulse noise	64.29%	60.88 %	0.5905	55.762	18.824	1856.8	28.780	18.279
10% impulse noise	47.24%	47.89 %	0.4231	59.097	26.376	2611.3	37.584	15.375
15% impulse noise	38.47%	38.80 %	0.3322	60.624	32.009	3179.7	42.886	13.720
20% impulse noise	30.03%	31.33 %	0.2740	61.621	36.611	3650.2	46.529	12.578
JPEG 60%	89.77%	78.08 %	0.9619	6.485	3.663	432.3	3.733	33.304
JPEG 40%	90.10%	77.11 %	0.9065	8.620	5.609	668.7	5.726	29.561
JPEG 20%	89.95%	76.82 %	0.8560	10.672	7.233	868.2	7.365	27.378
JPEG 10%	88.96%	76.14 %	0.7807	14.172	9.197	1164.3	9.330	25.153

Table 2. Correlation coefficient between the recognition accuracy and the image quality
metrics for all distortions used in experiments

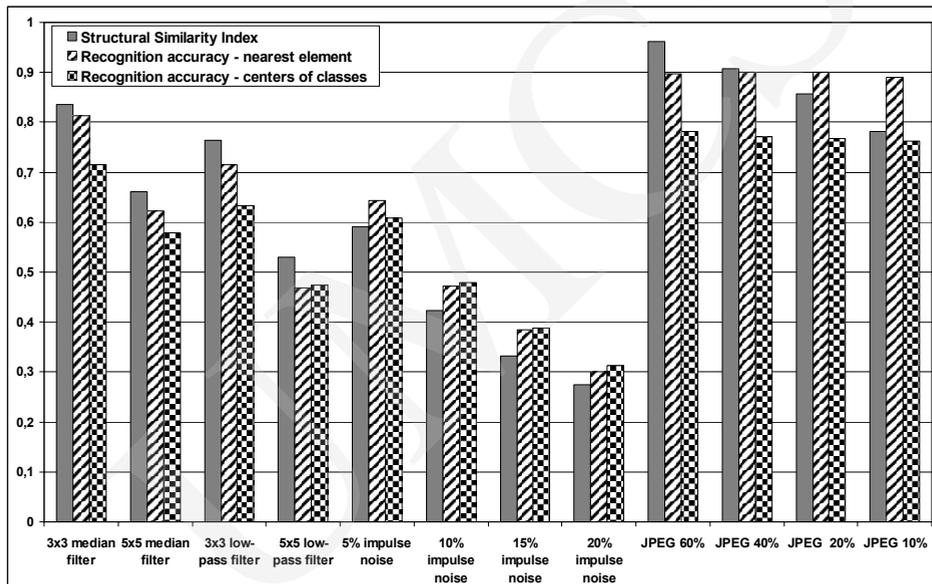| Classification approach	Correlation coefficient					
	SSIM	M_{SVD}	M_{DCT}	M_{DFT}	M_{DWT}	PSNR
Nearest element	**0.9709**	-0.8858	**-0.9592**	-0.8297	-0.9236	0.9218
Centers of classes	**0.9716**	-0.8738	**-0.9608**	-0.8227	-0.9258	0.9166

Fig. 4. Illustration of the dependence between the recognition accuracy and the SSIM index

Table 3. Correlation coefficient between the recognition accuracy and the image quality metrics
excluding the JPEG compressed images

Classification approach	Correlation coefficient					
	SSIM	M_{SVD}	M_{DCT}	M_{DFT}	M_{DWT}	PSNR
Nearest element	**0.9764**	-0.7200	-0.9313	-0.6706	-0.8826	0.8807
Centers of classes	**0.9668**	-0.7032	-0.9276	-0.6574	-0.8775	0.8642

Conclusions

Treating PSNR values as the basic ones because of the popularity of that
method of image quality assessment, regardless of its poor correlation with
Human Visual System, it can be noticed that two of the metrics (SSIM and
M_{DCT}) are much better correlated with the obtained recognition accuracy. An
interesting fact is relatively poor correlation of some other recently proposed
metrics, especially based on Singular Value Decomposition and DFT. However,

the metric based on Discrete Cosine Transform is noticeably less correlated with the obtained recognition results if the JPEG compressed images are excluded from the analysis.

Nevertheless, Structural Similarity seems to be a very interesting solution for predicting the texture recognition accuracy using the 2DLDA algorithm in the presence of image distortions. An interesting direction of our future research may be the verification of the presented approach (especially usefulness of SSIM index) for color images, not necessarily textures, also utilizing some other pattern recognition techniques.

References

[1] Brodatz P., *Textures: A Photographic Album for Artists and Designers*. Dover, New York, (1966).

[2] Swets D.L., Weng J., *Using Discriminant Eigenfeatures for Image Retrieval*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(8) (1996) 831.

[3] Tsapatsoulis N., Alexopoulos V., Kollias S., *A Vector Based Approximation of KLT and Its Application to Face Recognition*. Proceedings of the IX European Signal Processing Conference EUSIPCO-98, Island of Rhodes, Greece, September, 3 (1998) 1581.

[4] Swets D. L., Weng J., *Hierarchical Discriminant Analysis for Image Retrieval*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(5) (1999) 386.

[5] Jing Xiao-Yuan, Tang Yuan-Yan, Zhang D., *A Fourier-LDA Approach for Image Recognition*, Pattern Recognition, 38 (2005) 453.

[6] Kukharev G., Forczmański P., *Face Recognition by Means of Two-Dimensional Direct Linear Discriminant Analysis*. Proceedings of 8th International Conference on Pattern Recognition and Information Processing, Minsk, Belarus, May 18-20, (2005) 280.

[7] Kukharev G., Forczmański P.: *Facial Images Dimensionality Reduction and Recognition by Means of 2DKLT*, Machine Graphics and Vision, Accepted 11.01.2008, (2008).

[8] Fukunaga K., *Introduction to Statistical Pattern Recognition*, 2nd Ed., New York, Academic Press, (1990).

[9] Kukharev G, Kuźmiński A., *Biometric Techniques. Part I. The Methods of Face Recognition*. Szczecin (Poland), Pracownia Poligraficzna WI PS, (2003) 310, in Polish.

[10] Eskicioglu, A., Fisher, P., Chen, S., *Image Quality Measures and Their Performance*, IEEE Transactions on Communications, 43(12) (1995) 2959.

[11] Miyahara M., Kotani K., Algazi V.R., *Objective Picture Quality Scale (PQS) for Image Coding*, IEEE Transactions on Communications, 46(9) (1998) 1215.

[12] Wang, Z., Bovik, A., *A Universal Image Quality Index*, IEEE Signal Processing Letters, 9(3) (2002) 81.

[13] Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E., *Image Quality Assessment: From Error Measurement to Structural Similarity*. IEEE Transactions on Image Processing, 13(4) (2004) 600.

[14] Shnayderman, A., Gusev, A., Eskicioglu, A., *A Multidimensional Image Quality Measure Using Singular Value Decomposition*, Proc. SPIE Image Quality and Systems Performance, 5294(1) (2003) 82.

[15] Girshtel, E., Slobodyan, V., Weissman, J., Eskicioglu, A., *Comparison of Three Full–Reference Color Image Quality Measures*, 18-th IS&T/SPIE Annual Symposium on Electronic Imaging, Image Quality and System Performance, San Jose, CA, Proc. SPIE, 6059 (2006).

[16] Wang, Z., Sheikh, H., Bovik, A., *No-Reference Perceptual Quality Assessment of JPEG Compressed Images*, IEEE International Conference on Image Processing, September (2002) 477.

[17] Picard, R. W., Kabir, T., Liu, F., *Real-Time Recognition with the Entire Brodatz Texture Database*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, June (1993) 638.