ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

NO-REFERENCE IMAGE QUALITY ASSESSMENT ALGORITHM FOR CONTRAST-DISTORTED IMAGES BASED ON LOCAL STATISTICS FEATURES

Ismail T. Ahmed[1, 2] and Chen Soong Der[3]
[1]College of Computer Science and Information Technology, Universiti Tenaga Nasional, Malaysia
[2]College of Computer Sciences and Information Technology, University of Anbar, Anbar, Iraq
[3]College of Graduate Studies, Universiti Tenaga Nasional, Malaysia
E-Mail: esmaeel006@yahoo.com

ABSTRACT

Contrast change is a special type of image distortion; it is a very important for visual perception of image quality. Most No-Reference Image Quality Assessment (NR-IQA) metrics are designed for the quality assessment of images distorted by compression, noise and blurring. Few NR-IQA metrics exist for Contrast-Distorted Images (CDI). Existing approaches rely on global statistics to estimate contrast quality. The current No Reference Image Quality Assessment for Contrast-Distorted Images (NR-IQACDI) uses global statistics features. Room for improvement exists, especially for the assessment results using the image database called TID2013 which has poor correlation with Human Visual Perception (HVP); Pearson Correlation Coefficient (PLCC) < 0.7. In this work, instead of relying on the global statistics features, NRIQACDI is presented based on the hypothesis that image distortions may alter the local region statistics (Local patches features). Our experiments are conducted to assess the effect of using local patches features with natural scene statistics (NSS). The experiment results are based on K-fold cross validation with K range from (2 to 10). The statistical tests indicate that the performance using local statistical features are better than that of the NRIQACDI. The use of other statistical features and selection methods should be further investigated to increase the quality of prediction performance.

Keywords: contrast change, contrast-distorted images (CDI), no reference image quality assessment for contrast-distorted images (NR-IQACDI), local statistical features, natural scene statistics (NSS).

1. INTRODUCTION

Contrast is a fundamental attribute of images that plays an important role in human visual perception of image quality. Although there are many ways to enhance image contrast, but very few dedicated to automatic quality assessment of contrast changed images. Therefore, Objective IQA algorithms are preferable in order to analyze the images and predict the quality of contrast changed images. Depending on the availability of an “ideal quality” original image, objective IQAs are classified into Full Reference (FR), Reduce Reference (RR), and No Reference (NR) [1]. Figure-1 shows the general taxonomy of IQA/VQA [2].

Loss of contrast and visible details in an original image may be attributed to the limitation of the acquisition device and poor lighting condition [3]. Therefore, the acquired image is the original source and hence the perfect-quality reference is unavailable in this case. As such, No Reference Image Quality Assessment Algorithm (NRIQAs) can be used to optimize the settings of the Contrast Enhancement (CE) algorithms by providing the quality of the output contrast changed images.

Contrast change, however, is distinct from the other distortion types for the reason that an image processed by a proper histogram mapping can obviously improve image contrast and visual quality. It is worth to mention here that the related work performed in the area of NR-IQA for CDI is quite limited. Figure-2 illustrates the reference image with contrast distorted type [4].

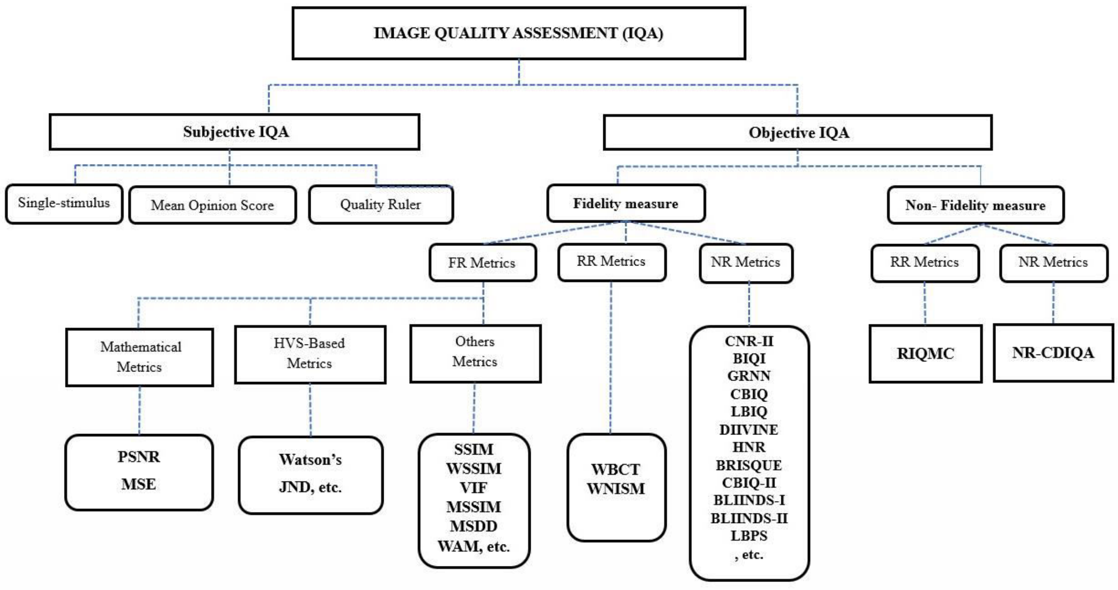

Figure-1. IQA measurement classifications [2].

(a) (b) (c) (d)

Figure-2. (a) Original 'lenna' image, (b) low contrast (dark) 'lenna' image, (c) low contrast (bright) 'lenna' image, (d) histogram equalized 'lenna' image [4].

However, most existing state-of-the-art IQA algorithms are dedicated for CDI based on global statistical features such as mean, variance, skewness, kurtosis, and entropy of pixel intensities [5, 6, 7]. In [5], a no-reference (NR) IQA method for contrast enhancement was proposed based on the principle of natural scene statistics (NSS). In [6], a reduced-reference (RR) method was proposed based on moment statistics. In [4], the contrast quality is determined by the histogram flatness and spread. The performance of the state-of-the-art NR-IQACDI still requires improvement, especially for the assessment results using the image database called TID2013 which has poor correlation with Human Visual Perception (HVP); Pearson Correlation Coefficient (PLCC) < 0.7. Therefore, in this work, instead of relying on the global statistics features, NRIQACDI is presented based on the hypothesis that image distortions may alter the local region statistics (Local patches features). Image processing using local patches has become very popular and was shown to be highly effective and useful. Patches are powerful primitives in the area of Image Processing. Our experiments are conducted to assess the effect of using local patches features with NSS. The experiment results are based on K-fold cross validation with K range from (2 to 10). The statistical tests indicate that the performance using local statistical features are better than that of the NRIQACDI. In the next section (Section 2), the proposed method stages including feature extraction, optimal Fitting Distribution Selection and image quality prediction. Section 3 describes the experimental results and Section 4 concludes the current work.

2. PROPOSED METHOD STAGES

Here, the proposed method consists of three steps and its block diagram is shown in Figure-3. We first decompose an image into local image patches 128x128. Then, for each image patch, five statistical features such as mean, standard deviation, skewness, kurtosis and entropy are extracted and build NSS models upon them using a large-scale image database. Each local statistical feature from all patches of the image are then pooled together by averaging the computed values to obtain final feature vector. Finally, by support vector regression (SVR), final features vector is mapped to the image's subjective quality score for NR IQA.

VOL. 13, NO. 17, SEPTEMBER 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

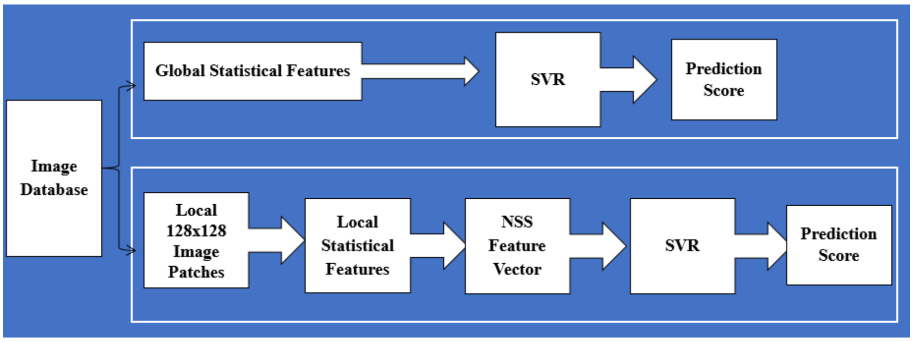

Figure-3. Proposed method block diagram.

2.1 Feature extraction step

For each image in the SUN2012 database [7], we first decompose an image I into local patch image of the size 128 x 128. Then, we compute local patches features for each input local 128 x128 image patches. Let μ denotes the sample mean operator. Then, for each image patch, five features such as sample mean me(Q), standard deviation std(Q), entropy ent(Q), kurtosis ku(Q), and skewness sk(Q) are computed as:

$$\mathrm{me}(Q) = \mu(Q), \tag{1}$$

$$\mathrm{std}(Q) = \sqrt{\mu[(Q-\mu(Q))^2]}\,, \tag{2}$$

$$\mathrm{sk}(Q) = \frac{\mu[(Q-\mu(Q))^3]}{\mathrm{std}(Q)^3}, \tag{3}$$

$$\mathrm{ku}(Q) = \frac{\mu[(Q-\mu(Q))^4]}{\mathrm{std}(Q)^4} - 3, \tag{4}$$

$$\mathrm{ent}(Q) = -\sum_j p_j(Q)\log_2 p_j(Q), \tag{5}$$

where $p(Q)$denotes the frequency of intensity value *j* occurs in Q. Each local statistical feature from all patches of the image are then pooled together by averaging the computed values to obtain final feature vector.

To build our NSS models based on these features, compute the probability distribution function for them based on the approach of NSS taken from SUN2012 database [7] which includes 16873 images that cover a large variety of image content. And show that the statistical features correlate with HV perception of contrast distortion.

2.2 Optimal fitting distribution selection step

In order to identify the best fit distribution, dfittool is used to fit the data. Trials for various distribution models were done and managed to fit with models such as normal, logistic and non-parametric.

The histograms of these features are shown in Figure-4. The histogram of the mean feature in Figure-4 (a) is found to be well fitted by a Gaussian probability density function. The fitting curve of mean is also plotted in Figure-4(a).

For the standard deviation and kurtosis features, their histograms, as given in Figures 4(b) and (c), respectively, can be well fitted by generalize extreme value probability density functions. The fitting curves of standard deviation and kurtosis features are also shown in Figures 4(b) and (c), respectively.

For the entropy and skewness features, their histograms, as given in Figures 4(d) and (e), respectively, can be well fitted by non-parametric probability density functions. The fitting curves of entropy and skewness features are also shown in Figures 4(d) and (e), respectively.

In order to compute NSS features of five features, apply the distribution to the data using the probability distribution function, by computing pdf values for the distribution *pd* at the values of five features *x* as follows:

$$Y = \mathrm{pdf}(pd, x) \tag{6}$$

2.3 Image quality prediction step

For any given image, the set of five local patches features is calculated with NSS models of mean, standard deviation, entropy, kurtosis, and skewness. The quality of the image is predicted based on this feature set. After the feature vector of a given image is calculated, we adopt Support Vector Regression (SVR) (via LIBSVM-3.12 package [8]) to determine the mapping function between the feature set and the subjective quality score. Then, a mapping is trained to predict the quality scores by using SVR. SVR is preferred because this machine learning algorithm has been successfully adopted in other NR-IQA approaches.

ARPN Journal of Engineering and Applied Sciences

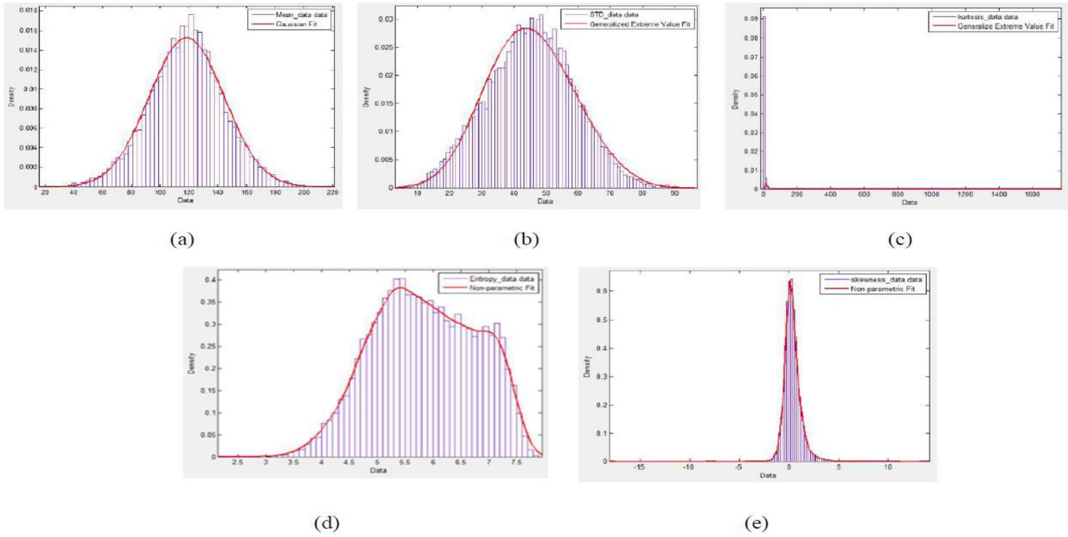

Figure-4. Histograms and the corresponding fitting curves of different features based on images in SUN2012 database [7] (a) Mean (b) Standard Deviation (c) Kurtosis (d) Entropy (e) Skewness.

3. EXPERIMENTAL RESULTS AND ANALYSIS

3.1 Images databases

For the experiments, we select the test images from the three publicly available databases, the CSIQ [9], TID2013 [10] and CID2013 [6]. We use only the contrast distorted images in the three databases (that is, reference images are excluded). A total of 116, 250, and 400 distorted images are selected from CSIQ, TID2013, and CID2013, respectively. The distorted image sizes for CSIQ, TID2013, and CID2013 were 512 x 512 pixels, 384 x 512 pixels, and 768 x 512 pixels, respectively. The difference Mean Opinion Scores (DMOS) associated with distorted images, which is ranging from 0 to 1, is reported, where a lower DMOS signifies a higher quality. We performed experiments on a laptop with an Intel (R) Core (TM) 2 Duo CPU, 2G RAM memory with a MATLAB R2013a platform.

3.2 Assessment of enhanced IQA performance

We use three performance metrics to assess the performance of IQA: The Spearman rank-order correlation coefficient (SROCC), Pearson's (linear) correlation coefficient (PLCC), and the root mean square error (RMSE) among the predicted objective scores and the subjective mean opinion scores (MOS). A good performance in terms of its correlation with human perception is normally indicated by SROCC, LCC~1, and RMSE~0. These metric measures the prediction monotonicity, prediction accuracy, and prediction consistency.

Given that regression is essentially a learning algorithm that requires training, K-fold cross validation (CV) was used for the assessment of the performance of IQA. The performance of IQA with and without the NSS features of local patches features was then assessed and to minimize bias using test images from 3 databases (CID2013, TID2013, and CSIQ) as well as k-fold cross validation with various k ranges from 2 to 10 followed by statistical tests. While performing K-fold cross validation, three databases are randomly partitioned into 10 subsets and 10-fold leave-one-out cross-validation is used to test the proposed metric. To reduce variability, multiple rounds of cross-validation (k = 2 to 10) are performed on different partitions. To avoid bias, the above cross-validation is repeated 100 times and the results are averaged as shown in Tables 1 and 2.

VOL. 13, NO. 17, SEPTEMBER 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Table-1. The average PLCC, SROCC and RMSE across 100 train-test rounds for three Databases using NRIQACDI [5] (five features).

k	CSIQ			TID2013			CID2013		
	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE
2	0.6296	0.6317	0.1317	0.4940	0.4652	0.8602	0.8390	0.8510	0.3409
3	0.6457	0.6274	0.1313	0.5165	0.4734	0.8456	0.8432	0.8536	0.3381
4	0.6444	0.6310	0.1303	0.5150	0.4774	0.8561	0.8422	0.8508	0.3359
5	0.6405	0.6105	0.1307	0.5165	0.4800	0.8458	0.8443	0.8588	0.3345
6	0.6397	0.6297	0.1308	0.5130	0.4725	0.8461	0.8448	0.8542	0.3350
7	0.6577	0.6109	0.1287	0.5163	0.4686	0.8469	0.8469	0.8502	0.3326
8	0.6633	0.6175	0.1267	0.5179	0.4765	0.8506	0.8460	0.8527	0.3336
9	0.6445	0.6087	0.1286	0.5267	0.4694	0.8452	0.8478	0.8554	0.3333
10	0.6517	0.6293	0.1284	0.5163	0.4826	0.8437	0.8457	0.8600	0.3339

Table-2. The average PLCC, SROCC and RMSE across 100 train-test rounds for three Databases using Local Patch Features with NSS (5 features).

k	CSIQ			TID2013			CID2013		
	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE
2	0.7641	0.7588	0.1116	0.6238	0.5105	0.7718	0.8380	0.8389	0.3408
3	0.8059	0.7793	0.1039	0.6293	0.5124	0.7649	0.8621	0.8604	0.3199
4	0.8203	0.7992	0.0998	0.6407	0.51127	0.7601	0.8534	0.8466	0.3255
5	0.8077	0.7804	0.1031	0.6378	0.53012	0.7584	0.8581	0.8494	0.3146
6	0.8325	0.8172	0.0979	0.6614	0.53626	0.7395	0.8636	0.8591	0.3167
7	0.8130	0.7844	0.1007	0.6326	0.49526	0.7602	0.8579	0.8523	0.3212
8	0.8235	0.8075	0.0999	0.6391	0.49712	0.7572	0.8585	0.8532	0.3243
9	0.8226	0.8042	0.0971	0.6361	0.49755	0.7541	0.8634	0.8597	0.3188
10	0.8213	0.7983	0.0982	0.6490	0.53930	0.7512	0.8558	0.8434	0.3140

Table-1 shows the average result of assessment using the feature vector obtained from spatial domain (NRIQACDI) [5]. Table-2 shows the average result of assessment using the feature vector obtained from local patches image with NSS. The result in Table-2 is better than the result in Table-1. The performance improved when the local was used. The next section discusses and identifies whether the differences in the performances among NRIQACDI [5] and NRIQACDI based on local statistical features are significant.

3.3 Statistical performance comparison

Let ci is performance metric values of NRIQACDI and cipatch is performance metric values of using local statistical features. Then we calculate the difference between the two-performance metrics for each k in each of the databases.

$$d_i = cipatch\ - c_i \qquad (7)$$

After that, the average of the percentage of differences over all the k and databases are computed. Percentage is measured by dividing the difference in performance over the absolute value of performance metric of ci.

$$dp = \frac{1}{n}(\sum_{i=1}^{n} d_i\ / abs(c_i)) \qquad (8)$$

where n is the number of all k in all databases. Absolute value is used to preserve the sign of difference of performance in the percentage (increment or decrement). Other possible patch image size such as 64 x64 and 256 x 256 can also be used here but in our experiments we found that the 128 x 128 provides a better accuracy. Table-3 shows the percentage of difference in each of the performance matrices.

In order to evaluate the statistical significance of performance of each metric, hypothesis testing based on the Paired T-tests is applied on the performance metric value obtained by NRIQACDI [5] and local statistical features in order to produce the p-value (see Table-4). In

VOL. 13, NO. 17, SEPTEMBER 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

general, p-value of < 0.05 shows that there is a significant difference within the values.

Table-3. Percentage Difference Results for NRIQACDI – Local Patches Features with NSS.

DB	64 x 64 Patch image			128 x 128 Patch image			256 x 256 Patch image		
	PLCC	**SROCC**	**RMSE**	**PLCC**	**SROCC**	**RMSE**	**PLCC**	**SROCC**	**RMSE**
TID2013	15.19%	3.57%	-6.71%	24.14%	8.52%	-10.76%	-30.60%	-33.28%	9.50%
CID2013	-22.10%	-30.95%	40.16%	1.46%	-0.30%	-4.03%	-19.36%	-23.69%	36.86%
CSIQ	11.93%	11.44%	-8.51%	25.67%	27.41%	-21.84%	23.25%	27.10%	-19.02%
Overall Databases	1.67%	-5.31%	8.30%	17.09%	11.87%	-12.21%	-8.90%	-9.95%	9.11%

Table-4. P-Values of differences between NRIQACDI and Local Patches Features with NSS.

DB	64 x 64 Patch image			128 x 128 Patch image			256 x 256 Patch image		
	PLCC	**SROCC**	**RMSE**	**PLCC**	**SROCC**	**RMSE**	**PLCC**	**SROCC**	**RMSE**
TID2013	significant	significant	significant	significant	significant	significant	Non	Non	Non
CID2013	Non	Non	Non	significant	Non	significant	Non	Non	Non
CSIQ	significant	significant	significant	significant	significant	significant	significant	significant	significant
Overall Databases	Non	Non	Non	significant	significant	significant	Non	Non	Non

3.4 Discussions

The aim of our experiments to assess the effect of using local patches features with NSS. Experiments results show the following:

- Table-3 shows that there is an improvement in the results of experiment using TID2013 database which is our primary target for improvement. There is a good increment in PLCC and SROCC by 24.14% and 8.52%, respectively. Meanwhile, there is a good decrement in RMSE by 10.76%. The three p-values for TID2013 shown in Table 4 are less than 0.05, indicating that there are significant differences in these three performance matrices.
- For CID2013 database, PLCC and SROCC decrements marginally by 1.46% and 0.30%, respectively. The RMSE increment by 4.03%. Nevertheless, the three p-values for CID2013 reported in Table-4 are less than 0.05, indicating that differences in these three performance matrices are significant statistically.
- For CSIQ database, there is a good increment in PLCC and SROCC by 25.67% and 27.41%, respectively. Meanwhile, there is a good decrement in RMSE by 21.84%. Also, the statistical tests indicate that there are significant differences in these three performance matrices.
- For the average results of the three databases, there is a good increment in PLCC and SROCC by 17.09% and 11.87%, respectively. On the other hand, there is a good decrement in RMSE by 12.21%. The three p-values shown in Table 4 are less than 0.05, indicating that the differences in these three performance matrices of all databases are statistically significant.

The statistical test shows that there are significant improvements in PLCC, SROCC and RMSE because $p<0.05$. Therefore, by using local Patches features with NSS, we have found that the performance of NRIQACDI can be improved.

4. CONCLUSIONS

In this research, instead of relying on the global statistics features, NRIQACDI is presented based on the hypothesis that image distortions may alter the local region statistics (Local patches features). Then could help to improve the performance of IQA in predicting image quality of contrast distorted images. In this study, instead of NR-IQA-CDI [5] relying on the global statistics features, NRIQACDI is presented based on the local statistics features. The experiments showed positive results that using local patches features with NSS could significantly improve the performance of IQA. The statistical tests indicate that the performance using local patches features with NSS are better than that of the NRIQACDI [5]. The use of other statistical features and selection methods should be further investigated to increase the quality of prediction performance.

REFERENCES

[1] ECE C. and Mullana M.M.U. 2011. Image quality assessment techniques pn spatial domain. Int. J. Comput. Sci. Technol. 2(3). DOI: 10.1.1.219.5535.

[2] ISMAIL T. AHMED, Chen Soong Der and BARAA TAREQ HAMMAD. 2017. Recent Approaches on No-Reference Image Quality Assessment for Contrast

Distortion Images with Multiscale Geometric Analysis Transforms: A Survey. Journal of Theoretical and Applied Information Technology. 95(3).

[3] Arici T., Dikbas S. and Altunbasak Y. 2009. A histogram modification framework and its application for image contrast enhancement. IEEE Transactions on image processing. 18(9): 1921-1935. DOI 10.1109/TIP.2009.2021548.

[4] Tripathi A.K., Mukhopadhyay S. and Dhara A.K. 2011. November. Performance metrics for image contrast. In Image Information Processing (ICIIP), 2011 International Conference on (pp. 1-4). IEEE. DOI: 10.1109/ICIIP.2011.6108900.

[5] Fang Y., Ma K., Wang Z., Lin W., Fang Z. and Zhai G. 2015. No-reference quality assessment of contrast-distorted images based on natural scene statistics. IEEE Signal Processing Letters. 22(7): 838-842. DOI: 10.1109/LSP.2014.2372333.

[6] Gu K., Zhai G., Yang X., Zhang W. and Liu M. 2013. September. Subjective and objective quality assessment for images with contrast change. In 2013 IEEE International Conference on Image Processing (pp. 383-387). IEEE. DOI: 10.1109/ICIP.2013.6738079.

[7] J. Xiao, J. Hays, K. Ehinger, A. Oliva, and A. Torralba. 2010. SUN database: Large-scale scene recognition from abbey to zoo. in IEEE Conf. Computer Vision and Pattern Recognition.

[8] Chang C.C. and Lin C.J. 2001. http:// www. csie. ntu. edu. tw/~ cjlin/ libsvm. LIBSVM: a library for support vector machines.

[9] Larson E.C. and Chandler D.M. 2010. Categorical image quality (CSIQ) database. Online, http://vision. okstate. edu/csiq.

[10] Ponomarenko N., Ieremeiev O., Lukin V., Egiazarian K., Jin L., Astola J., Vozel B., Chehdi K., Carli M., Battisti F. and Kuo C.C.J. 2013, June. Color image database TID2013: Peculiarities and preliminary results. In Visual Information Processing (EUVIP), 2013 4th European Workshop on (pp. 106-111). IEEE. Electronic ISBN: 978-82-93269-13-7.