 ISBN 978-3-200-02328-4

SIMPLIFIED STRUCTURAL SIMILARITY MEASURE FOR IMAGE QUALITY EVALUATION

Emil Dumic, Iva Bacic, Sonja Grgic

Department of Wireless Communications, University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia

ABSTRACT

In this paper we present the modified SSIM (SSIMmod) and simplified SSIM (SSIMsimpl) measures and compare them with 3 objective measures (MSE, original SSIM and VIF). Objective measures are compared using 7 publicly available subjective image quality databases and 3 correlation types (Pearson's, Spearman's and Kendall's correlation). Results show that SSIMmod gives higher correlation than original SSIM, while SSIMsimpl with similar weighted mean correlation as original SSIM achieves faster calculation time.

Index Terms**—** SSIM, Image quality database, VIF, correlation

1. INTRODUCTION

Image quality can be evaluated using different measures. The best way to do this is by making visual experiments, under controlled conditions, in which observers grade which image provides better quality. Such experiments are time consuming and costly. Much easier approach is to use some objective measure that evaluates the numerical error between the original image and tested one. In real world, there is no perfect way for objective assessment of image quality [1]. However, there is no current standard and objective definition of image quality.

Structural similarity index (SSIM) [2] is one of the typically used full-reference objective image quality measures. In this paper we will examine how it can be modified and simplified, without degradation in correlation with subjective grades (MOS, Mean Opinion Score or DMOS, Difference Mean Opinion Score).

This paper is organized as follows. Section II describes used subjective databases that were used for creating and testing modified and simplified SSIM measure. Section III describes MSE (Mean Squared Error), PSNR (Peak Signal to Noise Ratio), SSIM [2] and VIF [3] (Visual Information Fidelity) measures that were used for comparison with subjective grades. Here it is also explained SSIMmod and SSIMsimpl measures derived from original SSIM measure. Section IV explains performance measures used for comparing objective measures. Section V compares different objective image quality measures with results of subjective assessment and section VI gives the conclusion.

2. SUBJECTIVE IMAGE QUALITY DATABASES

7 different image quality databases were tested to determine correlation with objective measures:

- A57 (A57 database) [4]: 3 original images, 54 degraded images with 5 degradation types,
- CSIQ (Categorical Image Quality Database) [5]: 30 original images, 866 degraded images with 6 degradation types,
- LIVE (Laboratory for Image & Video Engineering) [6]: 29 original images, 779 degraded images with 5 degradation types,
- IVC (Image and video-communication) [7]: 10 original images, 185 degraded images with 5 degradation types,
- VCL@FER (Video Communications Laboratory @ FER) [8]: 23 original images, 552 degraded images with 4 degradation types,
- TID (Tampere Image Database 2008) [9]: 25 original images, 1700 degraded images with 17 degradation types,
- Toyama [10]: 14 original images, 168 degraded images with 2 degradation types.

Details about each of the above mentioned image quality databases can be found in references.

3. OBJECTIVE QUALITY MEASURES

3.1. MSE and PSNR measures

Mean Squared Error (MSE) measure is defined as:

$$\mathrm{MSE} = \frac{\sum_i \sum_j (a_{i,j} - b_{i,j})^2}{x \cdot y} \quad (1)$$

where in Eq. (1) a and b are original and distorted image. x and y are width and height of images.

Peak Signal to Noise Ratio (PSNR) measure is defined as:

$$\mathrm{PSNR} = 10\log_{10}\frac{255^2}{\mathrm{MSE}} \quad (2)$$

3.2. SSIM, SSIMmod and SSIMsimpl measures

Structural Similarity index (SSIM) is a novel method for measuring the similarity between two images [2]. It is computed from three image measurement comparisons: luminance, contrast and structure. At each step, the local statistics and SSIM index are calculated within the local window. Because resulting SSIM index map often exhibits undesirable "blocking" artifacts, each window is filtered with normalized Gaussian weighting function (11x11 pixels) prior calculation of the three components mentioned earlier. The Gaussian weighting function is described in Eq. (3) and afterwards normalized so that the sum of all filter values equals 1, Fig. 1. In [2] it is proposed that σ=1.5.

$$h(n_1,n_2) = \frac{1}{2\pi\sigma^2} e^{-\frac{n_1^2+n_2^2}{2\cdot\sigma^2}} \quad (3)$$

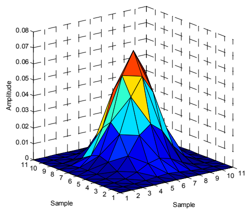

Fig. 1. Normalized Gaussian filter (σ=1.5)

In practice, one usually requires a single overall quality measure of the entire image, so mean SSIM index is computed to evaluate the overall image quality. The SSIM can be viewed as a quality measure of one of the images being compared, while the other image is regarded as of perfect quality. It can give results between 0 and 1, where 1 means excellent quality and 0 means poor quality. It is calculated over 11x11 pixels from three components, luminance, contrast and structure (after being filtered):

$$SSIM\_lum = \frac{2\cdot\mu_x\cdot\mu_y + C_1}{\mu_x^2+\mu_y^2+C_1} \quad (4)$$

$$SSIM\_cont = \frac{2\cdot\sigma_x\cdot\sigma_y + C_2}{\sigma_x^2+\sigma_y^2+C_2} \quad (5)$$

$$SSIM\_struct = \frac{\sigma_{xy}+\frac{C_2}{2}}{\sigma_x\cdot\sigma_y+\frac{C_2}{2}} \quad (6)$$

μ_x and μ_y are weighted mean values from original and degraded image blocks with size 11x11 pixels (N=11), using Gaussian weighting function from Eq. (3):

$$\mu_x = \sum_{i=1}^{N}\sum_{j=1}^{N} h(i,j)\cdot x(i,j)$$
$$\mu_y = \sum_{i=1}^{N}\sum_{j=1}^{N} h(i,j)\cdot y(i,j) \quad (7)$$

σ_x^2 and σ_y^2 are differences between filtered squared image and squared filtered image blocks (variances):

$$\sigma_x^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} h(i,j)\cdot x^2(i,j) - \mu_x^2$$
$$\sigma_y^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} h(i,j)\cdot y^2(i,j) - \mu_y^2 \quad (8)$$

σ_{xy} is similarly defined as difference between filtered product of original and degraded images and product of filtered original and degraded images (covariance):

$$\sigma_{xy} = \sum_{i=1}^{N}\sum_{j=1}^{N} h(i,j)\cdot x(i,j)\cdot y(i,j) - \mu_x\mu_y \quad (9)$$

C_1 and C_2 are constants defined as $C_1=(K_1L)^2$ and $C_2=(K_2L)^2$ where K_1 and K_2 are constants experimentally determined (K_1=0.01 and K_2=0.03). They help to improve stability of the measure when denominator is close to zero (for K_1 and K_2 equal to zero, this would be Universal Quality Index described in [11]). The final local SSIM measure is the product of Eqs. (4), (5) and (6):

$$SSIM = \frac{(2\cdot\mu_x\cdot\mu_y + C_1)(2\cdot\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \quad (10)$$

At the end, mean SSIM of all local SSIM values is calculated as their arithmetic mean.

In [12] it is suggested to downsample images prior calculating SSIM according to the:

$$F = \max(1, round(\min(M,N)/256)) \quad (11)$$

In Eq. (13) MxN is image size. Then we average images over FxF pixels and downsample images F times in horizontal and vertical direction. Afterwards, SSIM measure is calculated according to the Eq. (10).

In [13] it was suggested that luminance term is not relevant in computing SSIM measure. However, paper concludes that objective quality evaluation using Eq. (12) does not significantly alter the linear correlation between the DMOS and the objective ratings in comparison with original SSIM using Eq. (10). In the results section we will see that this is not correct, especially in image databases that have contrast degradations or mean shifts (like CSIQ and TID databases).

We will compute this SSIM measure, called SSIMmod, using only contrast and structure terms from Eqs. (5) and (6):

$$SSIM\ \mathrm{mod} = \frac{2 \cdot \sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (12)$$

However, when calculating variance and covariance terms (Eqs. (8) and (9)), weighted mean has to be also calculated, so that final SSIMmod measure has nearly equal calculation time when comparing with SSIM measure. In this paper we propose simplified SSIM calculation, SSIMsimpl. In SSIMsimpl, mean from original and degraded images are subtracted from images accordingly, after downsampling described in Eq. (11). Afterwards, σ_x^2, σ_y^2 and σ_{xy} in Eqs. (8) and (9) are calculated without μ_x and μ_y ($\mu_x = \mu_y = 0$). Also, here is σ=1 (for Gaussian weighting function) and K_2=0.06 because on average they yield somewhat better correlation results with subjective measures. However, using originally recommended parameters σ=1.5 and K_2=0.03 also produces similar correlation results.

3.3 VIF measure

The Visual Information Fidelity Criterion (VIF) [3] quantifies the Shannon information that is shared between the reference and the distorted images relative to the information contained in the reference image itself. It uses Natural Scene Statistics (NSS) modeling in concern with an image degradation model and an HVS model. Results of this measure can be generally between 0 and 1, where 1 means perfect quality and near 0 means poor quality, although it is possible to have grade higher than 1 (authors explain this as enhanced visual quality, e.g. contrast enhancement). In first step, original and degraded images are transformed using SPWT [14]. VIF is then calculated from 5 parameters, out of which 2 are calculated from error image and 2 from only original image. Fifth parameter, noise variance is experimentally determined (0.4 in later comparison).

4. PERFORMANCE MEASURES

4.1. Pearson's, Spearman's and Kendall's correlation

Each of the objective measures described earlier was graded using different performance measures: Pearson correlation coefficient, Spearman's rank correlation coefficient and Kendall's rank correlation coefficient.

Pearson's correlation coefficient is calculated according to:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1) \cdot s_x \cdot s_y}, \mathrm{i} = 1,..., n \quad (13)$$

In Eq. (13) x_i and y_i are grade values (x are objective grades and y are subjective measures), $\bar{x}$ and $\bar{y}$ are average grade values, and s_x and s_y are standard deviations, calculated by Eq. (14):

$$\begin{aligned} &\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i, \bar{y} = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i \\ &s_x = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n}(x_i - \bar{x})^2} \\ &s_y = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2} \end{aligned} \quad (14)$$

Because Pearson's correlation coefficient measures linear relationship between two variables, nonlinear regression should be done prior calculation of the correlation. The nonlinearity chosen for regression for each of the methods tested was a 5-parameter logistic function (a logistic function with an added linear term), as it was proposed in [15]:

$$Q(x) = b_1 \cdot \left(\frac{1}{2} - \frac{1}{1 + e^{b_2 \cdot (x - b_3)}}\right) + b_4 \cdot x + b_5 \quad (15)$$

However, this method has some drawbacks: firstly, logistic function and its coefficients will have direct influence on correlation (e.g. if someone chooses another function or even the same function with other parameters, results can be quite different). Another drawback is that function parameters are calculated after the calculation of the objective measures, which means that resulting parameters will be defined by the used image collection database. Different database can again produce different parameters. In [4] and [5] somewhat different logistic function is proposed, with 4 free parameters. Pearson's correlation was calculated using this function also, Eq. (16):

$$Q(x) = \frac{b_1 - b_2}{1 + e^{\frac{x - b_3}{b_4}}} + b_2 \quad (16)$$

We used three different methods to find the best fitting coefficients: Trust-Region method [16], Levenberg-Marquardt method [17] and Gauss-Newton method [18].

Final method for finding coefficients for nonlinear regression was the one which computed better results for performance measures (higher Pearson's correlation). An algorithm for optimizing coefficients b in (15) and (16) was developed. Firstly, set of 20 starting b parameters were checked to see which one gives best overall Pearson's correlation. For Eq. (15) $b_{1\text{-}5}$=[i, i, i, i, i] and [i, i+1, i+2, i+3, i+4], for $i \in \{1,10\}$; for Eq. (16) $b_{1\text{-}4}$=[i, i, i, i] and [i, i+1, i+2, i+3], for $i \in \{1,10\}$. Iterative algorithm for finding best b parameters was performed as long as difference between new and old Pearson's correlation was not under 0.0001. Best b coefficients were determined by the lowest RMSE (Root Mean Squared Error) after nonlinear regression, for every optimization method and every starting parameter. At the end, same iterative algorithm was performed, where starting parameters for every image database were chosen as ending (best) parameters of all other image databases (for the same image quality measure).

Spearman's correlation coefficient [19] is a measure of a monotone association that is used when the distribution of the data makes Pearson's correlation coefficient undesirable or misleading. Spearman's coefficient is not a measure of the linear relationship between two variables. It assesses how well an arbitrary monotonic function can describe the relationship between two variables, without making any assumptions about the frequency distribution of the variables. Spearman's correlation coefficient is calculated like Pearson's correlation in Eq. (13) over ranked variables. Rank of the sample in variable is its sorted location in a row. In the case of tied ranks, positions of all tied samples are calculated as an arithmetic mean of their ranks. If there are no any tied ranks, Spearman's correlation coefficient can be calculated simpler as:

$$\rho = 1 - \frac{6 \cdot \sum_{i=1}^{n} d_i^2}{n \cdot (n^2 - 1)} \qquad (17)$$

In Eq. (17) $d_i = x_i - y_i$ are differences between the ranks of each observation from the two variables being compared and n is the number of samples.

Kendall's rank correlation coefficient [20] is another performance measure which was used to compare objective and subjective measures. It measures the similarity of the orderings of the data when ranked by each of the quantities. All pairs of observations are ranked according to the first variable X (rank i) and then according to the second variable Y (rank j). Afterwards, every pair of observations from the first ranking is compared with all pairs of observations from the second ranking. Any pair of observations (x_i, y_i) and (x_j, y_j) are said to be concordant if the ranks for both elements agree: that is, if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. They are said to be discordant if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$ (case of tied ranks), the pair is neither concordant nor discordant. Final correlation coefficient is calculated as (τ_b coefficient):

$$\tau_b = \frac{n_{concordant} - n_{discordant}}{\sqrt{\left(\frac{N \cdot (N-1)}{2} - \sum_{i=1}^{T} \frac{t_i (t_i - 1)}{2}\right) \cdot \left(\frac{N \cdot (N-1)}{2} - \sum_{j=1}^{U} \frac{u_j (u_j - 1)}{2}\right)}} \qquad (18)$$

In Eq. (18) N is the number of observations, t_i is the number of t similar samples of variable X at rank $i \in \{1, \mathrm{T}\}$.

Similarly, u_j is the number of u similar samples of variable Y at rank $j \in \{1, \mathrm{U}\}$. In the case where there are no tied ranks, Kendall's correlation coefficient can be simplified and calculated as (τ_a coefficient), Eq. (19):

$$\tau_a = \frac{n_{concordant} - n_{discordant}}{\frac{N \cdot (N-1)}{2}} \qquad (19)$$

5. RESULTS

Pearson's (with 5 and 4 parameter function), Spearman's and Kendall's correlation are given in Tables 1, 2, 3 and 4 accordingly. When comparing images across multiple databases, correlation can be calculated as an arithmetic mean or weighted arithmetic mean, which is calculated as:

$$\text{Wtd\_mean} = \frac{\sum_{i=1}^{7} (w_i \cdot corr_i)}{\sum_{i=1}^{7} w_i} \qquad (20)$$

$$w_i = \{54, 866, 779, 185, 552, 1700, 168\}$$

In Eq. (20) w_i are database sizes (A57, CSIQ, LIVE, IVC, VCL@FER, TID and TOYAMA accordingly). Mean and weighted mean correlations are shown on Figs. 2, 3, 4 and 5. Pearson's correlation is calculated after nonlinear regression.

Calculation times for tested measures are given in Table 5. Computer configuration which was used for calculating all objective measures: Intel Q6600 @2400 MHz, 4 GB RAM, Windows 7 with Matlab program. Mean time was calculated for all images in TID database (512x384 pixels) and all degradation types and levels. Converting to grayscale and scaling images if needed were not taken in calculation time. VIF measure was calculated using [21] (with MEX files for cross correlation calculation).

Table 1. Pearson's correlation, using 5-parameter function for nonlinear regression

	MSE	SSIM	SSIMmod	SSIMsimpl	VIF
A57	0.69324	0.80188	0.80109	0.93523	0.69899
CSIQ	0.81536	0.86126	0.91525	0.85503	0.92775
LIVE	0.87305	0.94488	0.9447	0.92483	0.95983
IVC	0.72145	0.91194	0.91259	0.86381	0.90283
VCL@fer	0.8241	0.91436	0.9133	0.86812	0.89548
TID	0.58495	0.77317	0.80976	0.7896	0.80934
TOYAMA	0.64918	0.8887	0.88869	0.75672	0.91629
Mean	0.73733	0.87089	0.88362	0.85619	0.87293
Wtd_mean	0.72386	0.85092	0.87608	0.84105	0.87826

Table 2. Pearson's correlation, using 4-parameter function for nonlinear regression

	MSE	SSIM	SSIMmod	SSIMsimpl	VIF
A57	0.66947	0.80185	0.801	0.93382	0.61604
CSIQ	0.80299	0.85938	0.91236	0.85411	0.92526
LIVE	0.85823	0.93835	0.9383	0.91481	0.95924
IVC	0.72065	0.91165	0.91232	0.86361	0.90262
VCL@fer	0.81091	0.90892	0.90756	0.85679	0.89234
TID	0.56892	0.77153	0.80817	0.78693	0.80505
TOYAMA	0.62642	0.88771	0.8877	0.75389	0.91367
Mean	0.72251	0.86849	0.88106	0.852	0.85918
Wtd_mean	0.70945	0.84796	0.87292	0.8364	0.8744

Table 3. Spearman's correlation

	MSE	SSIM	SSIMmod	SSIMsimpl	VIF
A57	0.61763	0.80666	0.80662	0.90791	0.62228
CSIQ	0.8058	0.87563	0.92839	0.86832	0.91945
LIVE	0.87556	0.9479	0.94783	0.93068	0.96315
IVC	0.68844	0.90182	0.90285	0.85468	0.89637
VCL@fer	0.82465	0.91125	0.91001	0.87452	0.88665
TID	0.5531	0.77493	0.81771	0.80138	0.74907
TOYAMA	0.61319	0.87938	0.87943	0.75725	0.90767
Mean	0.7112	0.87108	0.88469	0.85639	0.84923
Wtd_mean	0.70611	0.85391	0.8813	0.84954	0.85068

Table 4. Kendall's correlation

	MSE	SSIM	SSIMmod	SSIMsimpl	VIF
A57	0.43007	0.60629	0.6049	0.75035	0.45944
CSIQ	0.60836	0.6907	0.75755	0.6834	0.75373
LIVE	0.68646	0.79629	0.79603	0.7698	0.82701
IVC	0.52175	0.72231	0.7262	0.68014	0.71581
VCL@fer	0.63614	0.73315	0.73108	0.68667	0.69244
TID	0.40275	0.57676	0.61347	0.61588	0.58605
TOYAMA	0.44428	0.69394	0.69394	0.55925	0.7315
Mean	0.53283	0.68849	0.70331	0.67793	0.68085
Wtd_mean	0.53248	0.67068	0.69847	0.66864	0.68671

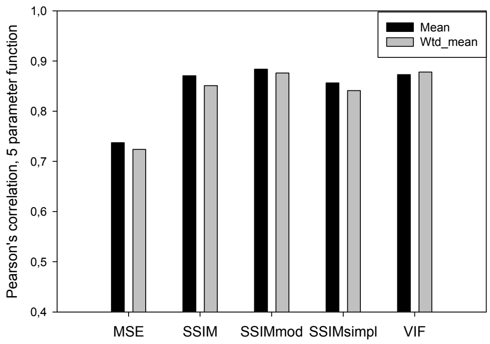

Fig. 2. Pearson's correlation, using 5-parameter function

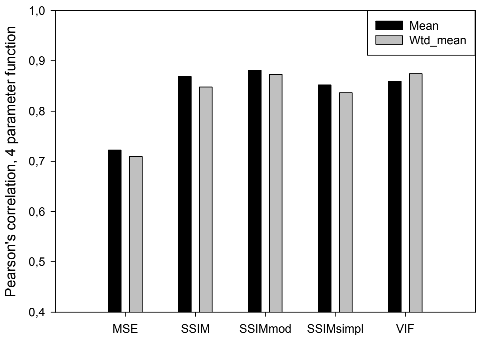

Fig. 3. Pearson's correlation, using 4-parameter function

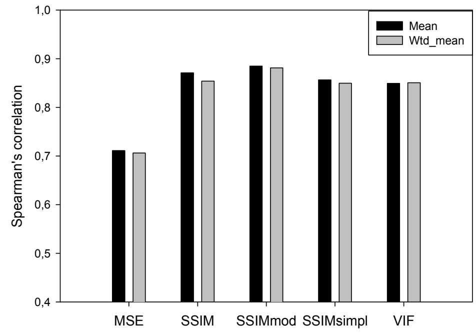

Fig. 4. Spearman's correlation

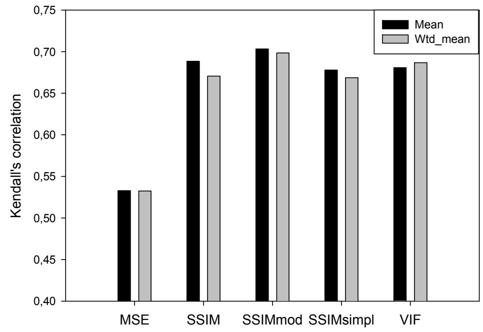

Fig. 5. Kendall's correlation

Table 5. Calculation time (t) and fps (frames per second) for tested objective measures

	MSE	SSIM	SSIMmod	SSIMsimpl	VIF
t (ms)	3.6	26.4	26.3	18.9	1322
fps (s^{-1})	277.8	37.88	38.02	52.91	0.756

From the correlation results it can be concluded that SSIMmod gives highest weighted mean Spearman's and Kendall's correlation for all tested objective measures. VIF measure gives highest weighted mean Pearson's correlation (for both 4 and 5 parameter fitting function), however SSIMmod has nearly the same weighted correlation while being faster about 50 times. When comparing Pearson's correlation results from 4 and 5 parameter fitting functions, it can be concluded that 4 parameter fitting function gives slightly worse results for all tested measures, however their rank stays the same. It can be also concluded that SSIMsimpl has similar weighted mean correlation (Pearson's, Spearman's and Kendall's) as original SSIM measure, with about 28% faster calculation time. Particularly, it has higher correlation than original SSIM in TID subjective image database which has 17 different degradation types. With 52.91 fps SSIMsimpl qualifies for real-time applications. In cases where somewhat lower speed is acceptable, SSIMmod (with only contrast and structure terms) should be used instead of the original SSIM measure. Luminance term should always be avoided because it can lower correlation in images with contrast degradations or mean shifts, while having same correlation results in databases with other degradation types.

6. CONCLUSION

In this paper we presented two modifications of the SSIM measure, SSIMmod and SSIMsimpl. SSIMmod achieves higher correlation results than original SSIM measure (while having the same calculation time) and SSIMsimpl with similar weighted average correlation outperforms SSIM with faster calculation time. In the future it is possible to develop standalone program for calculating tested objective measures, as well as some other newly developed objective measures.

ACKNOWLEDGEMENT

The work described in this paper was conducted under the research project "Picture Quality Management in Digital Video Broadcasting" (036-0361630-1635), supported by the Ministry of Science, Education and Sports of the Republic of Croatia.

REFERENCES

[1] Video Quality Experts Group, "Final Report from the Video Quality Experts Group on the Validation of Objective Models of Multimedia Quality", http://www.vqeg.org/, September 2008

[2] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[3] H.R. Sheikh.and A.C. Bovik, "Image information and visual quality," IEEE Transactions on Image Processing, vol.15, no.2, pp. 430- 444, Feb. 2006.

[4] http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html, access: 2011.

[5] http://vision.okstate.edu/index.php?loc=csiq, access: 2011.

[6] H.R. Sheikh, Z.Wang, L. Cormack, A.C. Bovik, "LIVE Image Quality Assessment Database Release 2", http://live.ece.utexas.edu/research/quality/subjective.htm, access: 2008.

[7] Patrick Le Callet, Florent Autrusseau: Subjective quality assessment IRCCyN/IVC database, http://www.irccyn.ec-nantes.fr/ivcdb/, access: 2011.

[8] A.Zaric, N.Tatalovic, N.Brajkovic, H.Hlevnjak, M.Loncaric, E.Dumic, S.Grgic, "VCL@FER Image Quality Assessment Database", Proceedings of the 53rd International Symposium ELMAR-2011, pp., 14-16 September 2011.

[9] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, "TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics", Advances of Modern Radioelectronics, vol. 10, pp. 30-45, 2009, http://www.ponomarenko.info/tid2008.htm, access: 2011.

[10] MICT (Media Information and Communication Laboratory) Image Quality Evaluation Database: http://mict.eng.u-toyama.ac.jp/mictdb.html, access: 2011.

[11] Z. Wang and A. C. Bovik, "A universal image quality index", IEEE Signal Processing Letters, Vol: 9 No: 3, pp. 81-84., 2002.

[12] https://ece.uwaterloo.ca/~z70wang/research/ssim/, access: 2011.

[13] D.M. Rouse and S.S. Hemami, "Understanding and simplifying the structural similarity metric", IEEE Int. Conf. on Image Process. (ICIP), pp. 1188-1191., 2008.

[14] E. P. Simoncelli, W. T. Freeman, "The Steerable Pyramid: A Flexible Architecture for Multi-Scale Derivative Computation", 2nd IEEE International Conference on Image Processing, Vol. 3, pp. 444-447., 1995.

[15] H.R. Sheikh, "Image Quality Assessment Using Natural Scene Statistics," Ph.D. dissertation, University of Texas at Austin, 2004.

[16] J.J. Moré and D.C. Sorensen, "Computing a Trust Region Step", SIAM Journal on Scientific and Statistical Computing, Vol. 3, pp. 553-572., 1983.

[17] D. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", SIAM J. Appl. Math. Vol. 11, pp. 431-441., 1963.

[18] J.E. Dennis, Jr., "Nonlinear Least-Squares," State of the Art in Numerical Analysis, ed. D. Jacobs, Academic Press, pp. 269-312., 1977.

[19] J. Hauke and T. Kossowski, "Comparison of values of Pearson's and Spearman's correlation coefficient on the same sets of data", Proceedings of the MAT TRIAD 2007 Conference, Bedlewo, Poland, 2007.

[20] M. Kendall, "A New Measure of Rank Correlation", Biometrika 30 (1-2), pp. 81-89., 1938.

[21] Visual Quality Assessment Package Version 1.1, http://foulard.ece.cornell.edu/gaubatz/metrix_mux/, access: 2011.