Distance transform measure based on edge-region information : An algorithm for image quality assessment

Zianou Ahmed Seghir[1], Fella Hachouf[2]
[1]Institute of sciences and technology, academic centre of Abbes Laghrour Khenchela
[2]Automatic and Robotic Laboratory, Mentouri Constantine University, Algeria,
e-mail: zianou_ahmed_seghir@yahoo.fr

Frédéric Morain-Nicolier[3]
[3]Centre de Recherche en STIC, IUT de Troyes, URCA, 9, rue de Québec, 10026 Troyes Cedex, France
e-mail : frederic.nicolier@univ-reims.fr

***Abstract*—in this paper we consider the problem of objective assessment of image quality. The Mean Structural Similarity Index (*MSSIM*) have been proposed to assess image quality, inspired by comparing the structures of the distorted and the reference images and presents image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion. We identify some limitations in this assessment method, and we propose an enhanced measure that corresponds more closely to visual judgment.**

Keywords- distance transform; quality assessment; distorted image

I. INTRODUCTION

The goal of an objective image quality [7, 15-16] assessment (IQA) is to develop quantitative metrics that can automatically predict perceived quality of image. IQA can be classified into subjective and objective. Objective image quality metrics serve primarily to assess the difference between two images, an original image and a distorted image. Subjective measurement (qualitative) is the most reliable judgment of the assessment of the images quality. It is carried out by the human observers. The working group includes expert observers and non-experts observers. A non-expert observer focuses its attention on the total sight, but a qualified observer can concentrate on the details.

Most methods that have been proposed for assessment of image quality in objective manner can be classified into three groups, the full-reference measure, the no-reference measure and the reduce-reference measure. The first class of methods addresses the full-reference methods, in which the algorithm has an access to a perfect version of the image to compare the deformed version. Hence, the no-reference methods do not required the reference image for IQA and has access only to the tested image and must assess the quality of the image without knowledge of the perfect version. The third class of image quality assessment is the reduced-reference method where the reference image is available partially; in some applications, one can transmit along with the compressed image a feature vector giving relevant information to control the quality of the result image. Methods based on these features are fast, but their relatively poor performances restrict their use to some specific applications.

The aim of this paper is to use distance transform (DT), with emphasis on gray-level DT. Particularly, the Weighted DT on curved space (WDTOCS) is used for IQA. So, in this paper, IQA is reduced to DT and region-edge information.

The rest of the paper is organized as follows. DT is presented in Section 2. Some algorithms of image quality measures are presented next. In Section 4, our proposed image quality measure is defined. Performance of the proposed method is compared with others measures using the images with different types of distortion in Section 5. We finish by the conclusion.

II. DISTANCE TRANSFORM

Distance is a fundamental concept in image analysis. The size and shape of an object can be used to detect or classify the object in a machine vision application, and distance information can be used in measuring both features. A distance transformation is operation, which determines the distance from picture element, or pixel, to a given subset pixels. The result, which is called a DT, or distance map, is an image, where the value of a pixel indicates its distance to the nearest feature or background pixel, depending on how the calculation area is defined. DT could be used in different domains: image retrieval, image classification, shape matching, image quality assessment etc… In literature there are many distance measures, using in image comparison [10]: Hausdorff distance [11] (HD), Baddeley's distance [9] and Euclidian distance etc... The HD has often been used in the content-based retrieval domain. It can be viewed as a dissimilarity measure between two binary images.

The DT on Curved Space (DTOCS) calculates distances along a gray-level surface, when gray-levels are understood as height values [6, 8]. Local distances, which are summed along digital paths to calculate the DT values, are defined using gray-level differences:

$$d(p_i, p_{i-1}) = |\mathcal{G}(p_i) - \mathcal{G}(p_{i-1})| + 1 \qquad (1)$$

where $\mathcal{G}(p)$ denotes the gray-value of pixel p, and p_i and p_{i-1} are subsequent pixels on a path. The Weighted distance transform on curved space (WDTOCS) is calculated from the height difference and the horizontal distance using Pythagoras theorem:

978-1-4577-1124-4/11/$26.00

$$d(p_i,p_{i-1}) = \begin{cases} \sqrt{|\mathcal{G}(p_i)-\mathcal{G}(p_{i-1})|^2+1}\,, & p_{i-1} \in N_4(p_i) \\ \sqrt{|\mathcal{G}(p_i)-\mathcal{G}(p_{i-1})|^2+2}\,, & p_{i-1} \in N_8(p_i)\backslash N_4(p_i) \end{cases} \quad (2)$$

The diagonal neighbors of pixel p are denoted by $N_8(p)\backslash N_4(p)$, where $N_8(p)$ consists of all pixel neighbors in a square grid, and $N_4(p)$ of square neighbors.
The optimal WDTOCS is defined as:

$$d(p_i,p_{i-1}) = \begin{cases} \sqrt{|\mathcal{G}(p_i)-\mathcal{G}(p_{i-1})|^2+a_{opt}^2}\,, & p_{i-1} \in N_4(p_i) \\ \sqrt{|\mathcal{G}(p_i)-\mathcal{G}(p_{i-1})|^2+b_{opt}^2}\,, & p_{i-1} \in N_8(p_i)\backslash N_4(p_i) \end{cases} \quad (3)$$

Where $a_{opt}^2 = (\sqrt{2\sqrt{2}-2}+1)/2 \approx 0.95509$
and $b_{opt}^2 = (\sqrt{2\sqrt{2}-2}+1)/2 \approx 1.36930.$

III. RELATED WORKS

In [3], a new philosophy of the Mean Structural Similarity Index (*MSSIM*) [3] is presented. However, it fails in measuring the badly blurred images [4]. In [4], an improved evaluation of quality assessment; called Gradient-based Structural Similarity (*GSSIM*) is proposed, based on the edge information as information of most significant image structure. This measurement is very interesting by employing the information of edge with *MSSIM* but it cannot be employed for the measurement quality of region information. In [5], Visual region of interest Weighted Quality Index (*VroiWQI*) is presented. This one is based on weighed indices of the local areas quality which capture the structural deformation in the local areas between the test image and the original image. *VroiWQI* is not interested to employ the edge information. In [2], an edge and region information with deformed Pixel (*ERDMSSIM*) is presented. This method is very interesting by using the edge and region information with the deformed pixel but it cannot be employed for the measurement quality of the displaced pixel. In [1], an improved *ERDMSSIM* is introduced by using the displaced pixel. This measure is called edge-region information with distorted and displaced pixels measure (*ERDM*). In this paper, the DT is used inside the proposed method (*ERDM)*, while distorted and displaced pixel is replaced by DT map. As results, edge-region information with DT measure (*ERTDM*) is developed.

IV. PROPOSED METHOD

To assess the quality of deformed image by a full-reference image quality metric, the complete description of the deformed image is compared with reference image. Several researches proved that the principal function of the human eyes is to extract the structure or edge information from the field of vision. So, HVS (the human visual system) is completely adapted for this purpose. The signals of the natural image are well structured; specifically the samples of the signals are strongly dependent between the others, especially when they are close in space. Consequently, in this research, one wants to develop a new measurement of the quality assessment based on the region-edge information and DT.

Before introducing the concept of the proposed measure, some useful concepts must be visited. The original image is represented by *Original* (*m, n*) and deformed image is noted by *Distorted* (*m, n*). The two images have *M* × *N* pixels. Thus, the images are partitioned into overlapped 11×11 blocks, where the overlapping area is on one pixel (see Fig.1). The choice of 11×11 overlap is motivated by the subdivision of the image on overlapped 8×8 blocks, often exhibits undesirable “blocking” artifacts problem [3]. Thus, this kind of problem is hidden, or mostly disappeared, when the partition superior on 8×8 block of pixel is used. Furthermore, the performance of developed measure is most excellent when 11×11 overlap is chosen.

In section 5, comparative study between the results of the proposed measure is made. For this purpose the performance of the measure is computed using three types of overlapping blocks, 8×8, 11×11 and 30×30.
The value of the total quality measurement is calculated in the overlapped blocks of 11×11 size, leading to a quality image map.

A. DT map by WDTOCS

In this part, DT of the original and deformed images haves previously been calculated with L.Ikonen et al. model [14] where DT function is *WDTOCS*. Then, the images are divided into overlapped 11×11 blocks. Hence, *DTd* and *DTo* are two blocks, belonging to the distance images: deformed and original image respectively.

In order to calculate the DT map, the following formula is used:

$$map\_dist(DTo, DTd) = \sqrt{\sum_{i=1}^{row}\sum_{j=1}^{col}|x(i,j)-y(i,j)|} \quad (4)$$

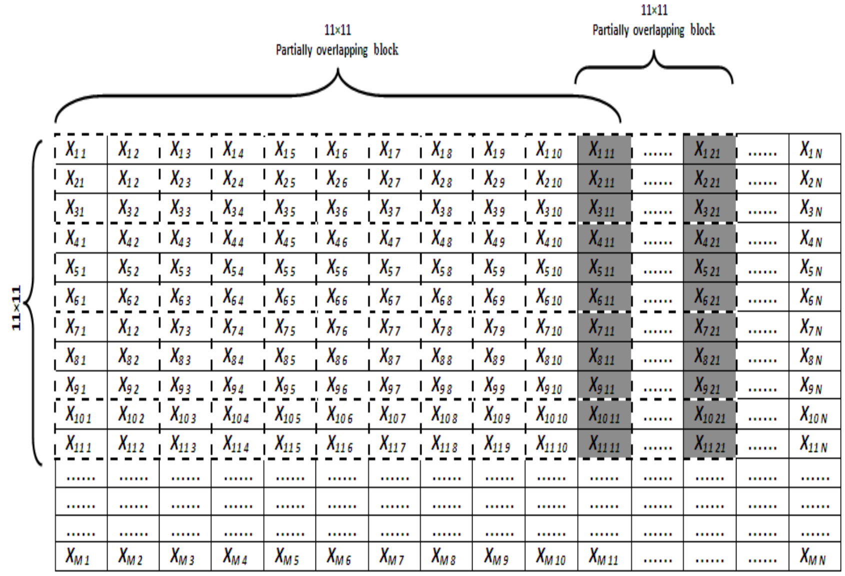
where *row* and *col* are blocks size (*DTd* , *DTo* respectively), $x(i,j)$ is a pixel of block *DTo* of original image and $y(i,j)$ is a pixel of *DTd* of deformed image.

B. Comparison measure of region structures

The areas in the original image are modeled by the entropy "***e***" in each block (11×11) of the original image using the following equation:

$$e = -\sum_{i=1}^{k} p(h_i)\log_2 p(h_i) \quad (5)$$

where h_i is a random variable indicating intensity, $p(h_i)$ is the histogram of the intensity levels in a region, K is the number of possible intensity levels. $K \in [0,256[$ for gray images.



Let db and ob be blocks of deformed and original images respectively, the formula calculating measurements of area structure comparison is as follows:

$$s_r(ob, db) = e * s(ob, db) \tag{6}$$

Where $s(ob,db)=(\sigma_{ob\,db}+C_3)*(\sigma_{ob}\sigma_{db}+C_3)^{-1}$; $C_3=C_2/2$, $C_2=(K_2L)^2$, $K_2 \ll 1$, L is the dynamic range of the pixel values (255 for 8-bit grayscale images).

Fig.2 shows the original image (Fig.2 a) and its normalized entropy map E (Fig.2 e), with a value of 1 indicating the region of highest interest represented as white, and a value of 0 indicating the region of lowest interest as black. The other levels of the image varying from white to black indicate the Visual regions with descending level of interest.

C. *Measure of edge information*

The comparison of contrast $c(ob, db)$ and the comparison of structure $s(ob, db)$ are respectively replaced by gradient of comparison contrasts $c_g(ob', db')$ and comparison of structure $s_g(ob', db')$.

$s_g(ob', db')$ is edge information measure:

$$s_g(ob', db') = \frac{\sigma_{ob'\,db'} + C_3}{\sigma_{ob'}\sigma_{db'} + C_3} \tag{7}$$

ob',db': are blocks (11×11) of original and deformed images which underwent a transformation by Sobel operator.

Fig.2. shows the edge image of original image (Fig.2 c) and edge image of distorted image (Fig.2 d).

D. *Edge-region information with DT measure (ERTDM)*

ERTDM map of original and deformed image is computed as sum of $ERTDM(ob, db)$:

$$ERTDM(ob, db) = [l(ob, db)]^{\alpha}.\left[c_g(ob', db')\right]^{\beta}.[s_{ER}(ob, db)]^{\gamma} + \omega * [1 + map\_dist(DTo, DTd)]^{-1} \tag{8}$$

Where: $s_{ER}(ob, db) = (e_{i,j} * s_g(ob', db') + s_r(ob, db))/2$, $c_g(ob', db') = (2\sigma_{ob'}\sigma_{db'} + C_2) * (\sigma_{ob'}^2 + \sigma_{db'}^2 + C_2)^{-1}$ and $\omega = e$. We set $\alpha=\beta=\gamma=1$.

Standardized measure is defined as:

$$ERTDM(Original, distorted) = \frac{\sum_{i=1}^{Eb} ERTDM_i(ob_i, db_i)}{2 * \sum_{i=1}^{Eb} e_i} \tag{9}$$

Where, Eb is the blocks number.

This measure is close to “*one*” when the image has the best quality and close to “*zero*” in the other case. In the proposed measure, the structure comparison $s_E(ob', db')$ is used to compute the related error to JPEG2000, JPEG and fast-fading Rayleigh channel model. Whereas $c(ob, db)$ is replaced in equation (8) by $c_g(ob', db')$ for White noise (WN) and Gaussian blur (Gblur).

The performance of the proposed measure is increased by using $c_g(ob', db')$ in calculation of WN and Gblur error and decreased in the other case. Thus, these indications are very important in the computation of *ERTDM.*

Furthermore, *ERTDM* is intended for color images.

In order to compute the distance map of image, the formula (4) is used, the difference between two images is the base of the formula. If the images (original and deformed) are the same, then formula turns the value “0” and if not it turns a value superior to “0”.

The distance image reflects a good difference than the original image; the sum of the subtraction between two distance images gives the error. So, the way that we see to compute the difference is the subtraction.

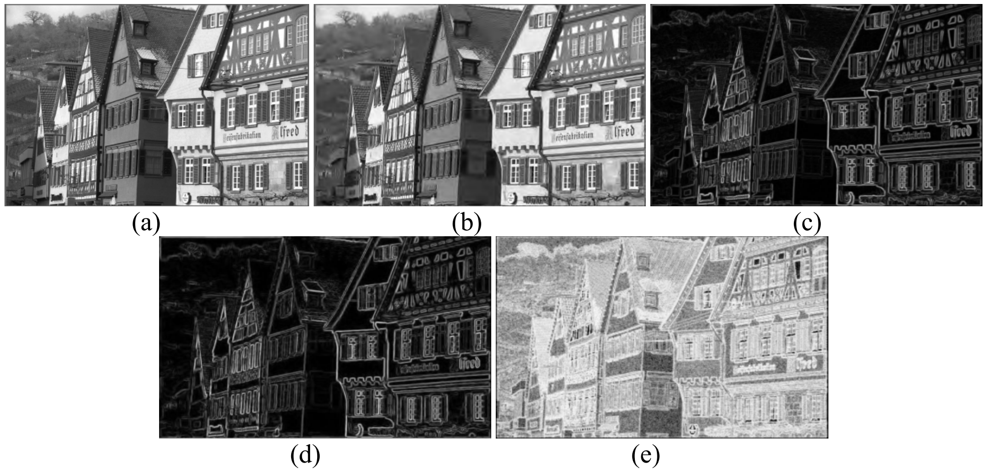


Figure 2. Edge and visual region of orginal and distorted images. (a) Original image. (b) distorted image. (c) edge image of original image. (d) edge image of distorted image. (e) and Visual regions image of original image.

The region information is then calculated by using the formula (6), and this information is extracted by the entropy operator. Moreover, Original and deformed images have to go through an edge detection operator. The gradient images are computed by applying the Sobel operator. The gradient images are then utilized in the formula (7); which determines edge information using the structure comparison. The formula (8) is the mixture of the previous formulas; its first part is the multiplication of luminance, contrasts, and the structure with the entropy operator and the second part is the distance map where it is standardized to take value "1". Lastly, the total error is computed by the formula (9).

V. RESULTS

The experiments are carried out on the set of images, to test the proposed measure. Five kind of distorted and compression methods are compared [13]: the JPEG2000 (227 images), the well-known JPEG (233 images) coding, White noise in the RGB components (174 images), Gaussian blur in the RGB components (174 images), and bit errors in JPEG2000 bitstream when transmitted over a simulated fast-fading Rayleigh channel (174 images).

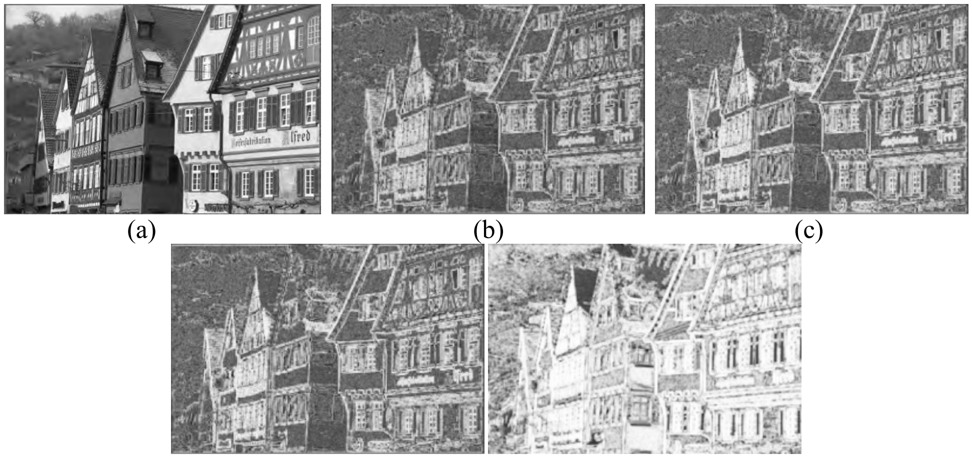
In order to provide quantitative measures on the performance of *ERTDM* quality assessment method, we follow the standard performance evaluation procedures employed in the video quality experts group (VQEG) FR-TV Phase II test [12], where mainly five evaluation metrics are used : The Pearson linear correlation coefficient (*CC*), The Spearman rank-order correlations coefficient (*ROCC*), the Root mean square prediction error (*RMS*), the Maximum absolute prediction error (*MAE*) and the Outlier ratio (*OR*). This metric assesses an objective model's ability to give consistently accurate predictions for all types of images and do not fail excessively for a subset of images, i.e., prediction consistency. The larger *CC* value means the better accuracy. The higher *ROCC* value means the better prediction monotonicity. While, the smaller *MAE*, *RMS* and *OR* values mean the better performance.

Table 1 summarizes the validation results.

Before comparing the proposed method with other measures, a comparative study is made to prove the choice of 11×11 size. It has been observed from Table 1, that *CC*, *MAE* and *RMS* are better when block size is 11×11. According to this study, the overall performance of the proposed technique reduces when block size is lower or greater than 11×11.

We have also compared the performance of *ERTDM* against *PSNR* (peak signal-to-noise-ratio), *VroiWQI*, *GSSIM, MSSIM, ERDM* and *ERDMSSIM.*

From Fig.3, the local dissimilarity map of *ERTDM* is clearly better than others. The Difference Mean Opinion Score (*DMOS*) against the objective quality measures (*ERTDM*, *PSNR, VroiWQI*, *GSSIM, MSSIM, ERDM* and *ERDMSSIM*) are represented by the scatter plots (see Fig.5). Looking at the curves (Fig.5), the *ERTDM* values are very closer to *DMOS*, proving the efficiency of this measure. However an interesting result is obtained from the comparison of the *ERTDM* with *PSNR, VroiWQI*, *GSSIM, MSSIM, ERDM* and *ERDMSSIM* in Table 1. The values of *CC* and *ROOC* are closer to 1; this means that *ERTDM* has a similar performance as the methods or earlier works. Furthermore, the performance of the *PSNR* is more less than previous one, for instance the *CC*=0.8709 and *ROOC*=0.8755. This one means that *PSNR* is not very well matched to perceived visual quality. The examination of the results obtained with *MSSIM*, lead us to say that *ERTDM* has a significant performance and this one is better than *MSSIM.* The blurred images (b) and (d) in Fig.4 almost have the same value of *PSNR* (31.4895 and 31.3934 respectively), but their visual quality are different, the subjective quality of blurred "lighthouse" image (Fig. 4 b) is much worse than the Blurred "building2" image (Fig.4 d), while the *MSSIM* values are contrary to the perceptual quality, the blurred image (Fig. 4 b) has a higher *MSSIM* value than the image Blurred "building2" image (Fig.5 d).



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

Figure 3. (a) Distorted image (JPEG2000 with *DMOS* =49.5811) and Images maps: (b) *ERTDM* (0.6864), (c) *ERDM* (0.6983), (d) *ERDMSSIM* (0.8203), (e) *MSSIM* (0.7558)

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

Figure 4. Comparison of "lighthouse" and "building2" images. (a) Original "lighthouse" image. (b) Blurred "lighthouse" image, *DMOS*= 84.489, *PSNR*= 31.4895, *MSSIM*= 0.5062, *ERTDM*= 0.3765. (c) Original "building2" image. (d) Blurred "building2" image, *DMOS*= 56.7104, *PSNR*= 31.3934, *MSSIM*= 0.448, *ERTDM*= 0.6424.

Let's see the Table 2. Moreover, a value of *DMOS* close to "100" means that the image has a bad visual quality, while a value close to "0" indicates that it has a good visual quality. The values of *MSSIM* and *PSNR* are less than *ERTDM*. The results clearly indicate that our *ERTDM* measure performs quite well and is competitive with other IQA measures.

Table1. Performance comparison of IQA methods (*ERTDM*, *PSNR, VroiWQI*, *GSSIM, MSSIM, ERDM* and *ERDMSSIM*) for JPEG, JPEG2000, Gaussian blurs, white noise, and Fastfading distorted images.

Model	***CC***	***ROCC***	***MAE***	***RMS***	***OR%***
PSNR	0.8709	0.8755	10.5248	13.4265	7.7597
VroiWQI	0.9330	0.9568	6.8620	8.3324	6.7821
GSSIM	0.9354	0.9584	5.8194	8.1736	6.8839
MSSIM	0.9367	0.9592	6.3997	8.0966	6.6497
ERDMSSIM	0.9405	0.9615	6.0770	7.8674	6.4664
ERDM	0.9557	0.9712	4.8981	6.8075	**5.7230**
*ERTDM(30*30)*	0.9314	0.9559	5.9334	8.415	6.4154
*ERTDM(8*8)*	0.9597	0.9737	4.6373	6.4997	6.0.387
*ERTDM(11*11)*	**0.9604**	**0.9740**	**4.5809**	**6.4439**	5.9470

Table 2. Performance comparison of IQA methods (*PSNR*, *MSSIM*, and *ERTDM*) on Gaussian blur distorted images.

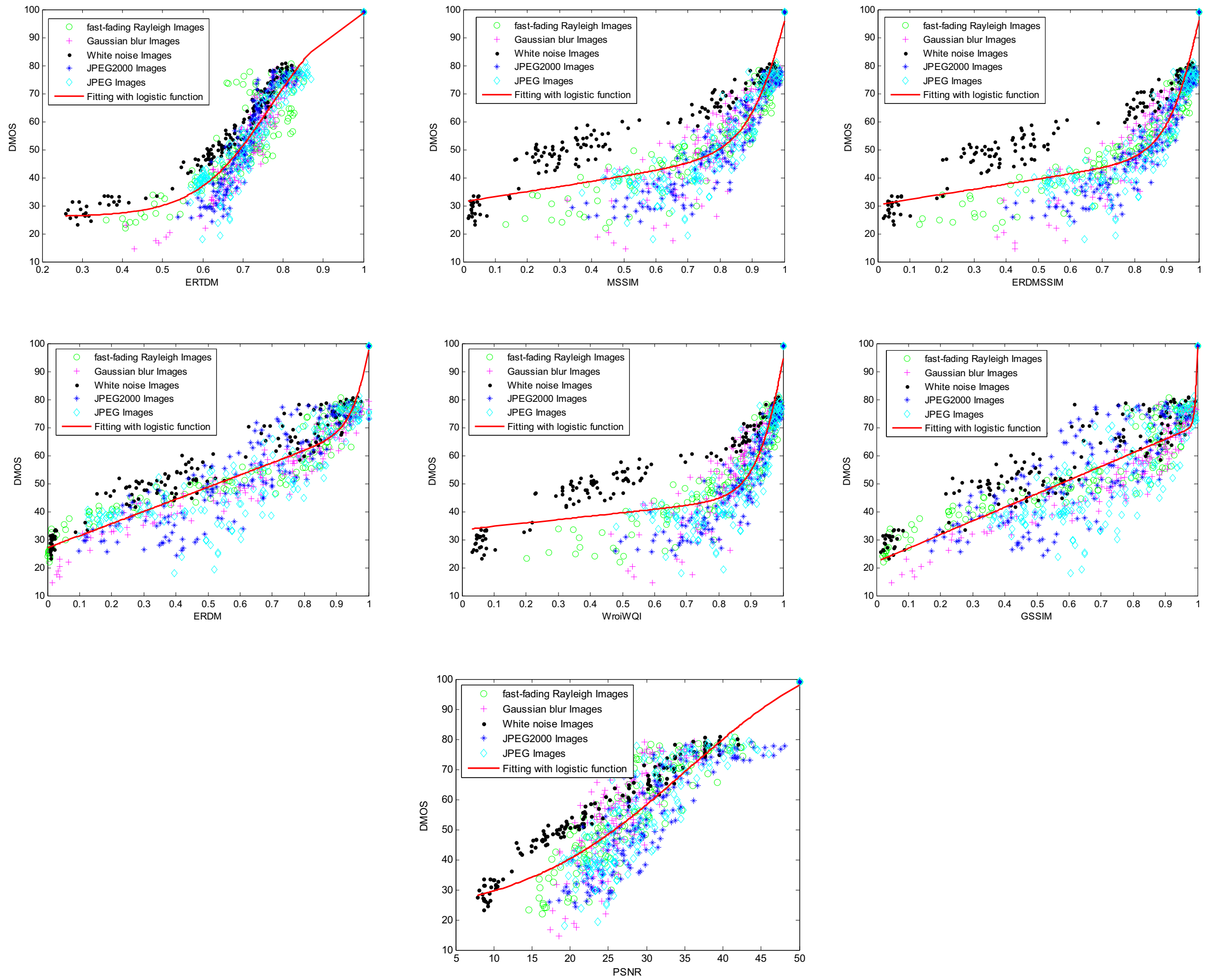
Model	***CC***	***ROCC***	***MAE***	***RMS***	***OR%***
PSNR	0.7834	0.7823	9.0550	11.4402	7.9310
MSSIM	0.9355	0.9882	5.6925	7.7146	6.6092
ERTDM	**0.9735**	**0.9951**	**3.6215**	**4.9706**	**5.5172**

VI. Conclusion

In this paper a DT measure is used for IQA. After, the transformations of the original and deformed images respectively to distance images using WDTOCS, the images are divided into overlapping blocks (11×11). The region-edge information is calculated using the structure similarity of images issue from Sobel and entropy operators.

Comparative study was done in this work. The obtained results were competitive with the previous works.

Future works following this study will include the use of a DT of real-images collection.



REFERENCES

[1] Z. Ahmed Seghir, F. Hachouf, “Edge-region information with distorted and displaced pixels measure for image quality evaluation,” in Proc. ISSPA 2010, pp. 77-80.

[2] Z. Ahmed Seghir, F. Hachouf, “Image Quality Assessment based on Edge-region information and Distorted pixel for JPEG and JPEG2000,” in Proc. ACIVS2009, pp. 156-166.

[3] Z.Wang A.C. Bovik, H. R. Sheikh, and E .P. Simocelli, “Image quality assessment: From error measurement to structural similarity,” IEEE Trans. Image Processing, vol. 13, no.4, pp.600- 612, Apr .2004.

[4] C. Guan-Hao, Y. Chun-Ling, X. Sheng-Li, “Gradient-based structural similarity for image quality assessment,” in Proc. ICIP06, pp. 2929-2932.

[5] D. Venkata Rao, N. Sudhakar, I. Ramesh Babu, L. Pratap Reddy,”Image quality assessment complemented with visual regions of interest,” in Proc. ICCTA 2007, pp. 681-687.

[6] P.J. Toivanen, “New geodesic distance transforms for gray- scale images,” Pattern Recognition Lett, vol. 17, n. 5, pp. 437-450, 1996.

[7] S.winkler, “Digital Video Quality: Vision Models and Metrics,“ Wiley Editorial Offices 2005.

[8] E. Baudrier, F. Nicolier, G. Millon and S. Ruan, “Binary-image comparison with local-dissimilarity quantification,” Pattern Recognition, vol. 41, no. 5, pp. 1461-1478, jan. 2008.

[9] D. Coquin et al., “Application of Baddeley's distance to dissimilarity measurement between gray scale images,“ PRL 2001.

[10] Di Gesù, V., Starovoitov, V., 1999, “Distance-based functions for image comparison,” Pattern Recoginition Lett, 1999.

[11] Huttenlocher, G.A. Klanderman and W.J. Rucklidge, “Comparing images using the Hausdorff distance,” PAMI 1993.

[12] Final VQEG report on the validation of objective quality metrics for video quality assessment: http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseI/

[13] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, “LIVE Image Quality Assessment Database Release2,” available at: http://live.ece.utexas.edu/research/quality

[14] L. Ikonen and P, “Toivanen. Shortest routes on varying height surfaces using gray-level distance transforms,” image Vision Comput, vol. 23, no. 2, 2005 , pp. 133–141.

[15] C. Yim and A.C. Bovik, “Quality assessment of de-blocked images,” IEEE Trans. Image Processing, vol. 20 , no. 1, January 2011 pp. 88-98.

[16] M.A. Saad, A.C. Bovik and C. Charrier, "A DCT Statistics-Based Blind Image Quality Index", *IEEE Signal Processing Letters*, pp. 583-586, vol. 17, no. 6, June 2010.