Color Distortion and Edge Feature for Perceptual Quality Assessment

Ahmed Zeggari [1], Zianou Ahmed Seghir [2], Mounir Hemam [2], Fella Hachouf [3] and Meriem Djezzar [4, 5]
E-mail : ahmed.zeggari@univ-tebessa.dz, zianou_ahmed_seghir@yahoo.fr, hemam.mounir@univ-khenchela.dz, hachouf.fella@gmail.com, meriem.djezzar@univ-khenchela.dz
[1] Math & Computer Sciences Dept. University of Tebessa, Tebessa. Algeria
[2] Khenchela University, ICOSI Laboratory, BP 1252 El Houria, 40004 Khenchela, Algeria
[3] Automatic and Robotic Laboratory, Mentouri Constantine University, Algeria
[4] University of Abbes Laghrour, Khenchela, Algeria
[5] LIRE Laboratory, Constantine 2 - Abdelhamid Mehri -University, Constantine, Algeria

The color distortion effect has an important impact on the perceived quality, which is ignored in previous related works. Unified with the color distortion outcome and edge similarity, a new full-reference image quality assessment was proposed named the gradient similarity-based distorted pixel and deformed color measure (GDCM). The components RGB of the color image are converted into image coded in YIQ color space. Then, Ruderman operators and the gradient images are calculated from the Y component. I and Q elements are used to identify the color distortion. Finally, the previous results are combined to compute the ultimate measure. Experimental results on databases illustrate that the GDCM performs very well.

Povzetek: Prispevek predstavlja novo tehniko za oceno kakovosti slike, ki temelji na barvnem popačenju in značilnostih robov.

1 Introduction

The image quality is reduced by different kinds of degradations. So, its quality needs to be assessed. Using the automatic image quality assessment (IQA) measures instead of the subjective one, is the purpose of the research in this area. As mentioned in the previous sentence, there are two ways to measure the image quality: the subjective methods are the first one, where the human observer is the ultimate judge of image quality. In the other hand, objective methods seek to assess the image quality by using machine. These objective IQA methods of identifying reliable quality are of three types, namely: full-reference (FR) measure, no-reference measure (NR) and reduced-reference (RR) measure. The perfection version of image is used to compare the distorted one in FR method [41-44]. NR [39-40] do not need a reference image and has only access to image test, its quality is evaluated without knowing the ideal version. RR methods assess the quality of test image using some features of reference image and the entire test one. In this work, FR IQA method is introduced.

Several researchers have already established an understanding of the human visual system (HVS) [25] in order to apply it for image quality. The most important ambition of an objective IQA measure, is to produce a method that imitates the HVS, then, the quality of a perceived image is instinctively evaluated. To reach this goal, it is significant to compare the measure performance with the subjective assessment. In the literature, it has been proved that the most important role of human eyes is to pick the contour or structure information from vision field. Hence, HVS is totally suitable for this goal. The issue of developing an objective assessment measure is the principle reason of this research work. This later can abolish the expensive study of the subjective assessment, where the scores values generated from it are more closer to human judgment. For an important use, a new method based on distorted pixel and deformed color (GDCM), which is useful for practical applications, such as compression and quality improvement, is developed to assess the image quality.

Experiments are carried out using sets of reference images with its deformed versions from TID2008 database [15], CSIQ database [16]; and LIVE database [14], TID2013 database [17]. Some existing models are involved in the comparison of the proposed method such as SSIM [9, 27], PSNR [9]; VSNR [9, 24], MS-SSIM [9, 1]; IFC [9, 22], NQM [9, 26]; VIF [9,23], DCTex [4]; Fsim [7], MAD [11]; GSM [10], GSDM [8]; and GVRO (the gradient similarity and visual region of interest based Ruderman operator measure and the distance transform (DT) [5] [20], DTSSIMC (color distortion and gradient similarity combined with structure, luminance and contrast comparison method) [21].

The rest of the paper is organized as follows. In section 2, some algorithms of the image quality

Table 1: Summary of the Related works.

Ref	Year	Pearson linear correlation coefficient (CC)										
		DATABASES										
		MICT [30] (2000)	LIVE JPEG/ JPEG2000 [14] (2003)	LIVE [14] (2004)	IVC [29] (2005)	LIVE [14] (2005)	A57 [31] (2007)	toyama [28] (2008)	CSIQ [16] (2009)	TID2008 [15] (2009)	WIQ [32] (2010)	TID2013 [17] (2013)
NQM [26]	2000	/	/	/	/	/	/	/	/	/	/	/
MS-SSIM [1]	2003	/	0.969	/	/	/	/	/	/	/	/	/
SSIM [27]	2004	/	/	0.967	/	/	/	/	/	/	/	/
IFC [22]	2005	/	/	/	/	0.929	/	/	/	/	/	/
VIF [23]	2006	/	/	/	/	0.950	/	/	/	/	/	/
VSNR [24]	2007	/	/	/	/	0.889	/	/	/	/	/	/
MAD [11]	2010	/	/	/	/	0.968	/	0.895	0.950	0.831	/	/
Fsim [7]	2011	0.908	/	/	0.938	0.960	0.925	/	0.912	0.874	/	/
DCTex [4]	2011	/	/	/	used	used	used	used	used	used	/	/
GSM [10]	2012	/	/	/	0.923	0.952	0.904	0.925	0.914	0.858	/	/
GSDM [8]	2015	/	/	/	0.925	0.961	0.903	/	0.955	0.879	0.87	/
DTSSIMC [21]	2018	/	/	/	/	0.953	0.932	/	0.947	0.887	/	0.889
GVRO [20]	2021	/	/	/	0.940	0.961	0.933	0.903	0.914	0.877	0.867	/

measurement are presented. GDCM is presented in section 3. Results are reported in Section 4. Section 5 is conclusions and future works.

2 Related works

The most used measures are MSE and PSNR. These later are failed in the measurement of visual quality as perceived by the human visual system (HVS). Therefore, numerous attempts have been made to design IQA measures inspired from the human visual system (HVS). These types of metrics highlight the importance of the sensitivity of the HVS to different visual signals, for instance contrast, frequency content, luminance and the interaction between different signal components. The two examples are the noise quality measure (NQM) [26] and the visual signal-to-noise ratio (VSNR) [24].

In the few past years Z. Wang et al. proposed a new metric named Structural Similarity Index (SSIM) [27] based on the measurement of structure distortion. In SSIM three components are calculated; the structure, contrast and luminance are its fundamental elements. Although its successful in IQA, SSIM struggle in the evaluation of the noisy and blurry images [2]. Several works are introduced in the literature [2, 3] trying to improve this latter by combining new features with the original one (SSIM). The first improvement is presented in [2] called GSSIM (Gradient Similarity). Sobel operator is used to achieve the edge information due to its simplicity and efficiency. This later is utilized in determination of edge information where the contrast comparison and structure comparison in [27] are replaced by the gradient-based contrast comparison and structure comparison respectively. This measure is very remarkable in employing edge information with SSIM [27], but it cannot be suitable for color distortion quality evaluation.

The multi-scale structural similarity (MS-SSIM) index [1] has resulted in much better performance than its single-scale counterpart. In [22, 23] information theory is presented for the judgment of image fidelity and two methods are proposed for the assessment of full-reference image quality. The information fidelity criterion (IFC) [22] quantifies the information shared between the reference image and the distorted images, while the visual information fidelity model (VIF) [23], which is a extension of the information fidelity metric (IFC), employed the reference image as well to calculate the relative fidelity of the information. IFC and VIF were derived from a statistical model for natural scenes, a model for image distortions, while VIF is inspired from a model of the human visual system. All of them are able to evaluate the quality of an image. MS-SSIM and SSIM share a common error when combining a single quality from a local quality map, all positions are considered of equal importance. In VIF, images are subdivided into separate sub-bands and these sub-bands may have different weights in the pooling step; however, within each sub-band, each position is always given the same value. The most apparent distortion algorithm (MAD) [11] uses two techniques. Thus, local luminance and a contrast masking check out high quality images. Changes

in the local dimensions of spatial-frequency components are used for low-quality images. In [4], DCTex metric is proposed with consideration of the texture masking effect and contrast sensitivity function. The metric is able to detect image distortion that is common in real-world applications. Another measure [7] that uses the phase congruency [6] is proposed to enrich IQA literature. Even it gives good results, it took more time to make an assessment. The pixel-wise gradient magnitude similarity deviation (GSM) [10] explores the image gradient feature to record image local structures. GSM computes the similarity among the gradients of reference and distorted images, and then calculates a few more detail. The gradient similarity and distorted pixel measures (GSDM) are utilized for IQA in [8]. Following modifications of the reference and the distorted pictures with distorted pixel measure and gradient mask, the gradient similarity and distorted pixel measures map among the reference and the distorted images are calculated using a simple function. The distance transform, color distortion and gradient similarity combined with structure, luminance and contrast comparison method (DTSSIMC) [21] is proposed to deal with SSIM to process color pictures by adding the color distortion idea and DT [5]. The pictures were first turned to gradient images. The gradient pictures are then used to compute the traditional SSIM components. Following that, DT images are computed. The color distortion is finally established. Even, DTSSIMC employs color distortions, it took more time to predict image quality score. A novel FR IQA method called gradient similarity and visual region of interest based Ruderman [12] operator measure (GVRO) is suggested [20]. As a result, the proposed scheme increases the effectiveness of IQA measures by using edge information as well as regions of interest and the Ruderman measure. Five imaging components are calculated based on the image characteristic. Firstly, the reference and test pictures are converted using the Ruderman operator, yielding Ruderman images. Secondly, the visual region of interest is calculated using local entropy. Then, Canny and suggested gradient are utilized to obtain gradient pictures from reference and test pictures, as well as, gradient maps are created using prior images. Lastly, the proposed approach is derived from all of the measurements.

On the other hand, because structural information cannot represent color changes between the reference picture and the distorted image, the above IQA approaches cannot appropriately evaluate image color distortion. Furthermore, color images are increasingly being employed in our daily lives and, in addition, color information carries significant signals for visual perception. Color information should thus be considered in order to increase the accuracy of IQA approaches.

This work developed the color similarity (CS) method by proposing an enhancement of gradient similarity based distorted pixel measure (GSDM) [8] with the incorporation of color distortion component inside it. It is mentioned even the gradient operator extract the edge from image, and then it is used in IQA as the previous work [33-38], one found that the gradient operator cannot differentiate color distortion in image. To avoid these issues, the transformation of color image to another image coded YIQ color space [13] is investigated. Consequently, the gradient similarity based distorted pixel and deformed color measure (GDCM) is emerged.

To summary all related works used in this paper, the table 1 is established. The accuracy of these works is introduced by using the pearson linear correlation coefficient (CC), where the values close to value one or minus one are suitable for the best CC. Only the values of CC of database used in the related work are mentioned. Some values of CC are not introduced, the word "used" is utilized to show that this database is explored by this work but the CC is not computed, for instance, DCTex [4] work. The details on the formula of CC and databases are explained in the result section.
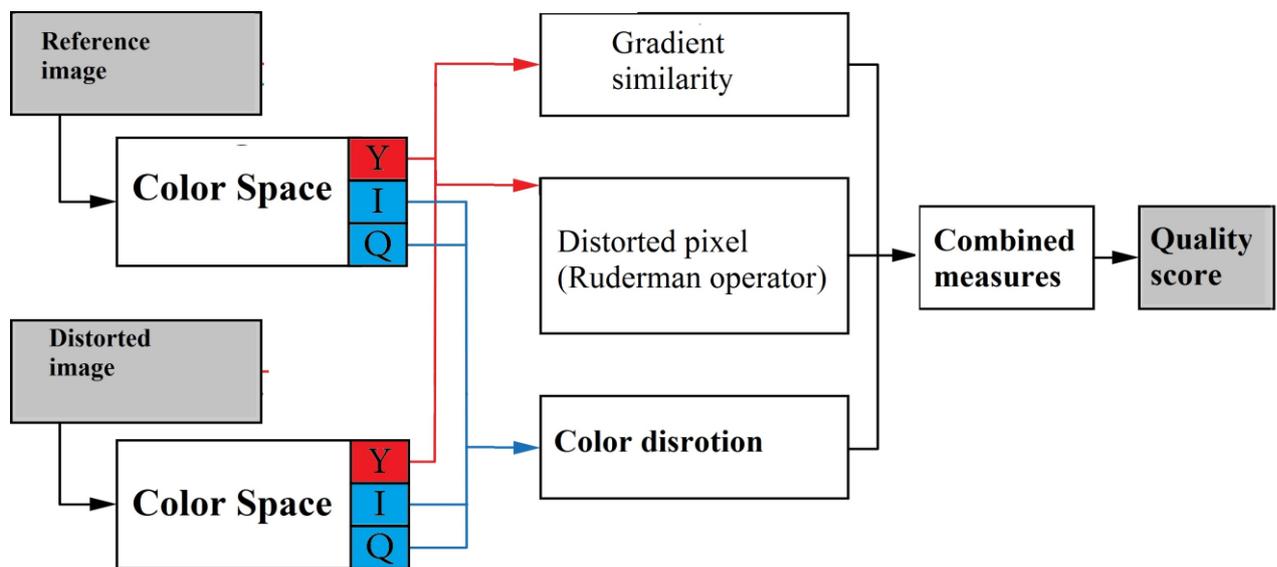

Figure 1: The proposed IQA measure framework.

3 The gradient similarity based distorted pixel and deformed color measure

In this stage, IQA is reduced to full-reference. It is known that HVS extracts the edge or structure from vision field. A new measure based on edge information is proposed taking into account this recent concept. Moreover, the pixel distortion is computed using Ruderman operator [12]. The color distortion is computed in this later to improve the former one [8]. The deformed and reference images are noted as $Dis(R, C)$ and $Ref(R, C)$ respectively. All of them have $R \times C$ pixels.

The proposed method structure is set as follows for this purpose:

- Compute edge information utilizing gradient similarity.
- Determine distorted pixels using Ruderman operator.
- Find color distortion.
- Determine spatial pooling approach for proposed measure.

A simple flowchart depicting computation of the proposed measure is shown in Figure 1.

The following variables and abbreviations used in the proposed method and the rest of manuscript are defined as:

Ref: reference image.
Dis: deformed image.
$R \times C$: image size.
CFI_map: color similarity map of hue (I).
CFQ_map: color similarity map of saturation (Q) channel information.
$DRef$: distorted reference.
$DDis$: deformed distorted image
Y: brightness.
$\mu(k, l)$: local mean within the 3×3 block surrounding position (k, l).
$\sigma(k, l)$: standard deviation of $Y(k, l)$ within the block.
DM_map: distorted map.
G_1 : gradients magnitude of reference image.
G_2 : gradients magnitude of deformed image.
G_map : *Gradient* map.
GDC_map: deformed color measure map.
GSDM: gradient similarity and distorted pixel measures.
ROCC: Spearman's rank order correlation coefficient.
SSIM: structural similarity index measure.
TID: Tampere image database.
CSIQ: Categorical Subjective Image Quality.
PSNR: peak signal-to-noise ratio.
CC: Pearson's linear correlation coefficient.
MS-SSIM: multi-scale structural similarity index measure.
NR: no-reference image quality assessment.
RR: reduced-reference image quality assessment.
FR: full-reference image quality assessment.
IQA: image quality assessment.
MOS: mean opinion score.
ROCC: Kendall's rank order correlation coefficient.

HVS : human visual system.
GVRO: gradient similarity and visual region of interest based Ruderman operator measure.
DT :the distance transform.
DTSSIMC: color distortion and gradient similarity combined with structure, luminance and contrast comparison method .
NQM: noise quality measure.
VSNR: visual signal-to-noise ratio.
IFC : information fidelity criterion.
VIF : visual information fidelity model.
MAD: most apparent distortion algorithm.
GSM: The pixel-wise gradient magnitude similarity deviation.
CS :Solor Similarity.
GDCM: Gradient Similarity based distorted pixel and deformed Color Measure.
*DMOS*P: predicted Difference Mean Opinion Score.
RMSE: Root mean square prediction error.
KROCC: Kendall rank-order correlation coefficient.

3.1 Color distortion measure

The edge information delivered by the gradient operator cannot show the distortion in color images. So, some particular measurements are used to deal with information delivered by color. The exchange between RGB to YIQ color space [13] is given as follows:

$$
\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)
$$

The chromatic channels of reference (*Ref)* and distorted (Dis) images are: I_1 (I_2), and Q_1 (Q_2), they measured with the equation (1). In addition, Y_1 and Y_2 are the gray images of *Ref* and *Dis* respectively. They will be used in computing the others metrics.

The chromatic features of I and Q are employed to calculate *CFI_map* and *CFQ_map*, the similarity is defined as follows:

$$
CFI\_map = \frac{2I_1 . I_2 + T_3}{I_1{}^2 + I_2{}^2 + T_3}
$$

$$
CFQ\_map = \frac{2Q_1 . Q_2 + T_3}{Q_1{}^2 + Q_2{}^2 + T_3} \quad (2)
$$

where $T_3 = (T_2 . 255)^2$, $T_2 << 1$.

3.2 Ruderman operator

The dissimilarity between the test and reference images is calculated by Ruderman operator [12]. The formulas that calculate the distorted reference ($DRef$) image is given by:

$$
\widehat{DRef}(k, l) = \frac{Y_1(k, l) - \mu(k, l)}{\sigma(k, l) + 1}
$$

$$
(3)
$$

Where, $k \in 1, 2 \ldots R$, $l \in 1, 2 \ldots C$ are spatial indices

$$\mu(k,l) = \frac{1}{(2II+1)(2JJ+1)} \sum_{ii=-II}^{II} \sum_{jj=-JJ}^{JJ} Y_1(k+ii, l+jj)$$

(4)

$$\sigma(k,l)$$
$$= \frac{1}{(2II+1)(2JJ+1)} \sqrt{\sum_{ii=-II}^{II} \sum_{jj=-JJ}^{JJ} [Y_1(k+i, l+j) - \mu(k,l)]^2}$$

(5)

In the previous formulas the local mean $\mu(k,l)$ is computed within 3×3 block. Also, the standard deviation $\sigma(k,l)$ of $Y_1(k,l)$ is measured with the same size of block. *II* and *JJ* values are fixed to 1. The deformed distorted image, which calculated with the same previous formulas, is named *DDis*.

Combining *DRef* and *DDIS* the distorted map (*DM_map*) is emerged and is expressed in the following formula

$$DM\_map = \frac{2DRef.DDis + T_1}{DRef^2 + DDis^2 + T_1}$$

(6)

where $T_1 = (T_2 . 255)^2$, $T_2 \ll 1$.

3.3 Edge similarity detection

The gradient process enables identify the highest change in the intensity or color in an image. The image's edges always seem to have a high gradient value. When there is a smooth area of the image, the gradient value is reduced. In the discrete domain, the gradient amplitude could generally computed using certain operators and estimate the derivative of the picture function using the difference between neighboring pixels. The alteration of each pixel in the picture is generally examined in the conventional image gradient method, and the first or second derivative of nearby edges is utilized to determine the gradient operator in the original image. For convolution computations, tiny region patterns are often employed. Edge image has important rule in vision field. This important information is produced from image using more edges detection. The gradient is employed in the proposed measure to compute this information. A new operator is proposed to generate this later. It is a pair of 3×3 convolution kernels, and it identifies the horizontal and vertical boundaries in images. It is introduced in the following formula:

	Gx	*Gy*
Mask	$\begin{pmatrix} 27.5 & 0 & -27.5 \\ 34 & 0 & -34 \\ 27.5 & 0 & -27.5 \end{pmatrix}$	$\begin{pmatrix} 27.5 & 34 & 27.5 \\ 0 & 0 & 0 \\ -27.5 & -34 & -27.5 \end{pmatrix}$

Gradient magnitude is given by equation (7)

$$|G| = \sqrt{Gx^2 + Gy^2}$$

(7)

Gradients magnitude for reference (Y_1) and deformed (Y_2) images are calculated using the previous equation. G_1 and G_2 are the gradients magnitude of reference and deformed images respectively. Afterwards, the *Gradient* map (G_map) is defined as

$$G\_map = \frac{2G_1.G_2 + T_1}{G_1{}^2 + G_2{}^2 + T_1}$$

(8)

3.4 Combined measures

In this stage all previous computed measures are combined to form the gradient similarity based distorted pixel and deformed color measure map (***GDC*_map**). Then, *GDCM* is computed from ***GDC*_map** as its standard deviation, it is represented by the following equations.

$$GDC\_map = DM\_map . G\_map.CFI\_map.CFQ\_map$$

(9)

$$GDCM = \sqrt{\frac{1}{R.C} \sum_{p=1}^{C} \sum_{q=1}^{R} (GDC - GDC\_map(p,q))^2}$$

(10)

Where

$$GDC = \frac{1}{R.C} \sum_{p=1}^{C} \sum_{q=1}^{R} GDC\_map(p,q)$$

(11)

Figure 2 shows a flowchart of High-level overview of the proposed method demonstrating the suggested measure's calculation.

4 Experimental results and analysis

This section contains the outcomes of our experiments. Section 4.1 first introduces the benchmark IQA datasets that were employed. Section 4.2 then discusses the assessment measures and procedure used. Section 4.3 concludes with a comparison of proposed method to the state-of-the-art.

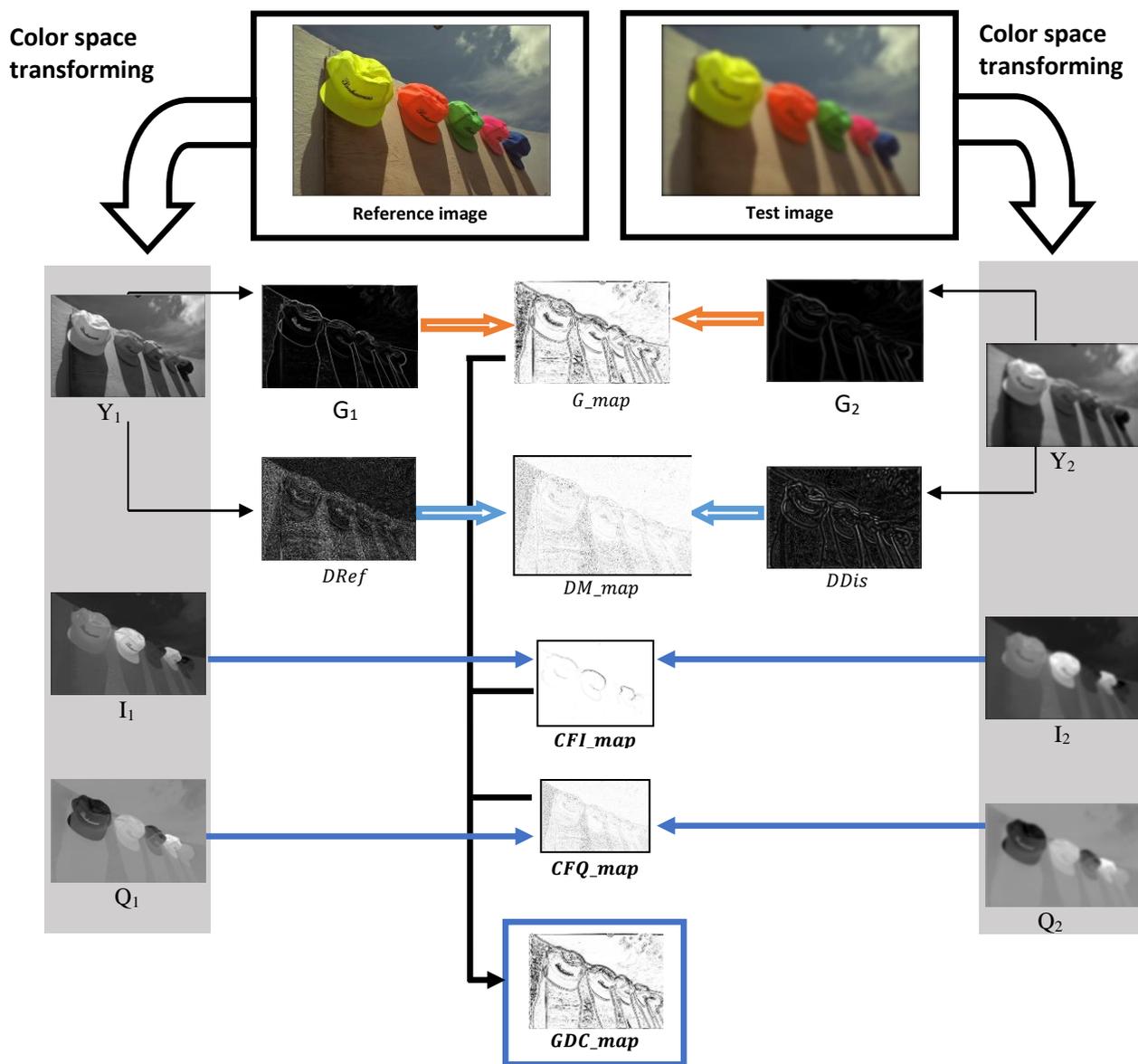Figure 3: Reference images in databases TID2008 and TID2013.

Figure 2: High-level overview of the proposed method.

4.1 Databases

The proposed method has been tested on many databases. The live database [14], TID2008 database [15], CISQ database [16] and TID2013 database [17] are used in these tests. In terms of the number of reference images, the number of distorted pictures, the number of quality distortion categories, the number of human observers, and the image format, these four databases are highly varied.

The Tampere Image Database (TID2008) [15] is intended for evaluation of full-reference image visual quality assessment metrics. TID2008 allows estimating how a given metric corresponds to mean human perception. The TID2008 contains 25 reference images (see figure 3) and 1700 distorted images (25 reference images × 17 types of distortions × 4 levels of distortions).

Reference images are obtained by cropping from Kodak Lossless True Color Image Suite. All images are saved in database in Bitmap format without any compression. File names are organized in such a manner that they indicate a number of the reference image, then a number of distortion's type, and, finally, a number of distortion's level: "iXX_YY_Z.bmp". The MOS was obtained from the results of 838 experiments carried out by observers from three countries: Finland, Italy, and Ukraine (251 experiments have been carried out in Finland, 150 in Italy, and 437 in Ukraine). Totally, the 838 observers have performed 256428 comparisons of visual quality of distorted images or 512856 evaluations of relative visual quality in image pairs. Higher value of MOS (0 - minimal, 9 - maximal, MSE of each score is 0.019) corresponds to higher visual quality of the image.

The Tampere Image Database (TID2013) [17] is the largest image quality database available in the public domain, both in terms of test images and number of subjects. TID2013 is an extension of TID2008. TID2013 is intended for evaluation of full-reference image visual quality assessment metrics. TID2013 allows estimating how a given metric corresponds to mean human perception. The TID2013 contains 25 reference images (see figure 3) and 3000 distorted images (25 reference images x 24 types of distortions x 5 levels of distortions). Reference images are obtained by cropping from Kodak Lossless True Color Image Suite.

The Categorical Subjective Image Quality (CSIQ) is an image database [16]. It comprises 30 original images (see figure 4) from which degraded images are generated using six types of degradations at four to five levels. The degradations used in CSIQ are: JPEG, JPEG2000, global contrast decrements, additive pink Gaussian noise, additive white Gaussian noise, and Gaussian blurring. CSIQ consists of 866 degraded images being the result of the deformation of the original images by the different types of previous degradations. CSIQ images are subjectively evaluated based on linear displacement of the images. Four liquid crystal display (LCD) screens at 1920×1200 resolution were calibrated following RGB color space and placed side by side with equal viewing distance to the viewer. Observers were instructed to keep the viewing distance stable at approximately 80 cm. All degraded versions of an original image were viewed simultaneously through the screen. Each observer placed these images in the screen, in which the horizontal distance between two images reflects the perceived

Figure 4: Categorical Subjective Image Quality (CSIQ) Image database: All 30 Source images.

quality of these images. An image placed close to the others means that: the observer thinks that their quality is not much different. On the other hand, placing a distant image to the left of the others means that the viewer thinks the old image is much worse in quality comparing to the last image. As a final step, image ratings are realigned. In this experiment, 35 observers participated. Observers include male and female with normal visual acuity or correct to normal. The ages of the observers range from 21 to 35. In general, the image database contains 5000 subjective estimates that are reported in the form of DMOS (Difference Mean Opinion Score).

The LIVE database [14] consists of 779 distorted images generated from 29 reference images (see figure 5). This later was constructed from twenty nine high resolution color images that were collected from the internet and CDs. These images consist of images of faces, people, animals, scenes, synthetic objects etc... Most of the images have the size 768×512 pixels. Five types of deformations were chosen in order to degrade the images. These types can occur in real applications. The types of deformation are: compression in JPEG format (233 images), compression in JPEG2000 format (227 images), Additive white Gaussian noise (AWN) in the RGB components (174 images), Gaussian blur (GB) in RGB components (174 images), and simulated fast fading Rayleigh channel (FF) (174 images). These distortions reflect a wide range of image degradations, from smoothing, to structured distortion, and random noise. The level of distortion was changed to produce images at a wide range of quality, from imperceptible levels to high levels of degradation that would significantly impede cognitive understanding of image content.

Table 2 summarizes the properties of these four databases.

4.2   Evaluation metrics and protocol

The non-linear mapping is computed among the objective and subjective scores [18]. A set of the predicted Difference Mean Opinion Score ($DMOS_P$) is obtained from the value calculated by the objective quality methods. Five parameters non-linear mapping (C_1, C_2, C_3, C_4 and C_5) are used in this process. This transformation is applied using a logistic function of the equation (12) [19].

$$DMOS_p = C_1 \text{fun}\big(C_2, (OM - C_3)\big) + C_4 + C_5 \qquad (12)$$

$$\text{fun}(h, OM) = \frac{1}{2} - \frac{1}{1 + \exp(OM\, h)} \qquad (13)$$

Where *OM* is the value calculated by the objective measure. C_1, C_2, C_3, C_4, C_5 are chosen for the best fit.

Then, four metrics are computed to benchmark the proposed method: the Spearman rank-order correlations coefficient (*ROCC*), the Pearson linear correlation coefficient (*CC*); and the Root mean square prediction error (*RMSE*)**,** Kendall rank-order correlation coefficient (*KROCC*).

The first index is the Spearman rank-order correlations coefficient (*ROCC*) [10], commonly

Figure 5: Reference images in databases Live.

Table 2: Four databases and their characteristics.

Database	Source Images	Distorted Images	Distortion Types	Image Type	Observers
TID2008	25	1700	17	color	838
CSIQ	30	866	6	color	35
LIVE	29	779	5	color	161
TID2013	25	3000	24	color	985

employed. The term correlation is used to designate the connection or the relation between any two variables. In statistics, this term is utilized to quantify the link between two quantitative variables. This junction can be either symmetric (possibility of permuting the quantitative variables x and y) or asymmetric (one of the values depends on the other, and therefore impossible to permute the quantitative variables x and y). To perform a Spearman rank correlation measurement between two quantitative variables (one representing the metric prediction values, and the other the MOS/DMOS values), the observations must be independent, the link between the two variables must be linear, and the distribution of the values of the variables must not follow a normal law. This method has the particularity of not using the values of the observations but their ranks. In the context of metrics without reference, the calculation proceeds as follows (the first column represents the predictions of the metric, the second column represents the values of the MOS or DMOS):
- Sorting of the first column (keeping the association with the values of the second column).
- Assign an ascending order number (rank) for the first sorted column.

- Sorting of the second column (keeping the association with the last rows).
- Assign an ascending number for the second sorted column.
- Calculate the square of the difference between each two rows.
Spearman's rank correlation coefficient is then calculated by:

$$ROCC = 1 - \frac{6 \sum (DMOS(i) - DMOS_p(i))^2}{n(n^2 - 1)} \tag{14}$$

Where the index i is the square of the difference of each two rows and n is the total number of observations.

The second one is the Pearson linear correlation coefficient (*CC*) between subjective (*DMOS*) and objective (*DMOS*$_P$) scores. It supplies an evaluation of the prediction accuracy and it is defined by:

$$CC = \frac{\sum_{i=1}^{n}(DMOS(i) - \overline{DMOS})(DMOS_p(i) - \overline{DMOS_p})}{\sqrt{\sum(DMOS(i) - \overline{DMOS})^2} \sqrt{\sum(DMOS_p(i) - \overline{DMOS_p})^2}} \tag{15}$$

Where the index *i* denotes the image sample and *n* denotes the number of samples.

The third one is the Root mean square prediction error (*RMSE*) between subjective (*DMOS*) and objective (*DMOS*$_P$) scores. It is defined by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(DMOS(i) - DMOS_p(i) \right)^2} \tag{16}$$

Finally, the Kendall rank-order correlation coefficient (KROCC) is defined as

$$KROCC = \frac{C - D}{0.5(N - 1)N} \tag{17}$$

where C stands for the number of concordant pairs between subjective (*DMOS*) and objective (*DMOS*$_P$) scores, while D denotes the number of discordant pairs and N is the number of observations.

These measures evaluate an objective aptitude model to provide consistently perfect predictions for every kind of images and do not fail extremely for a subset of images, i.e., prediction consistency. *ROCC* and *KROCC* evaluate the prediction monotonicity. *RMSE* and *CC* judge the prediction accuracy. Values close to one (1) or minus one (-1) are suitable for the best *ROCC, KROCC* and *CC*. While a big number is a sign of bad *RMSE*.

4.3 Discussion and performance comparison

Fourteen state-of-the-art measures are compared with GDCM to show the robustness of this latter. MS-SSIM, VSNR, PSNR, SSIM, NQM, DCTex, VIF, GSM, MAD, IFC, GVRO, DTSSIMC, Fsim and *GSDM* are the state-of-the-art methods used in this comparison.

The four metrics CC, KROCC, ROCC, and RMSE are implicated in the measurement of the performance of 14 methods and the proposed one. Tables 3–7 show all these results. The top three measures are highlighted in bold.

From the results presented in Tables 3–7, the proposed GDCM method is able to do better than other state-of-the-art algorithms in terms of CC and ROCC

values. Specifically, GDCM outperforms the first best method in terms of direct average ROCC values and the second best by approximately 0.001 with regard to direct average CC values (see Table 7).

However, an interesting result is obtained from the comparison of the GDCM with VIF, MS-SSIM; PSNR, Fsim; GSDM, and DTSSIMC in Tables 5. The values of CC and ROOC are closer to 1; this means that GDCM performs well against these methods or earlier works. In addition, the performance of VIF, MS-SSIM, PSNR, Fsim, GSDM and DTSSIMC in Table 5 is more less than GDCM, for instance VIF values are CC=0.928/ROOC=0.920, MS-SSIM values are CC=0.899/ROOC=0.913, PSNR values are
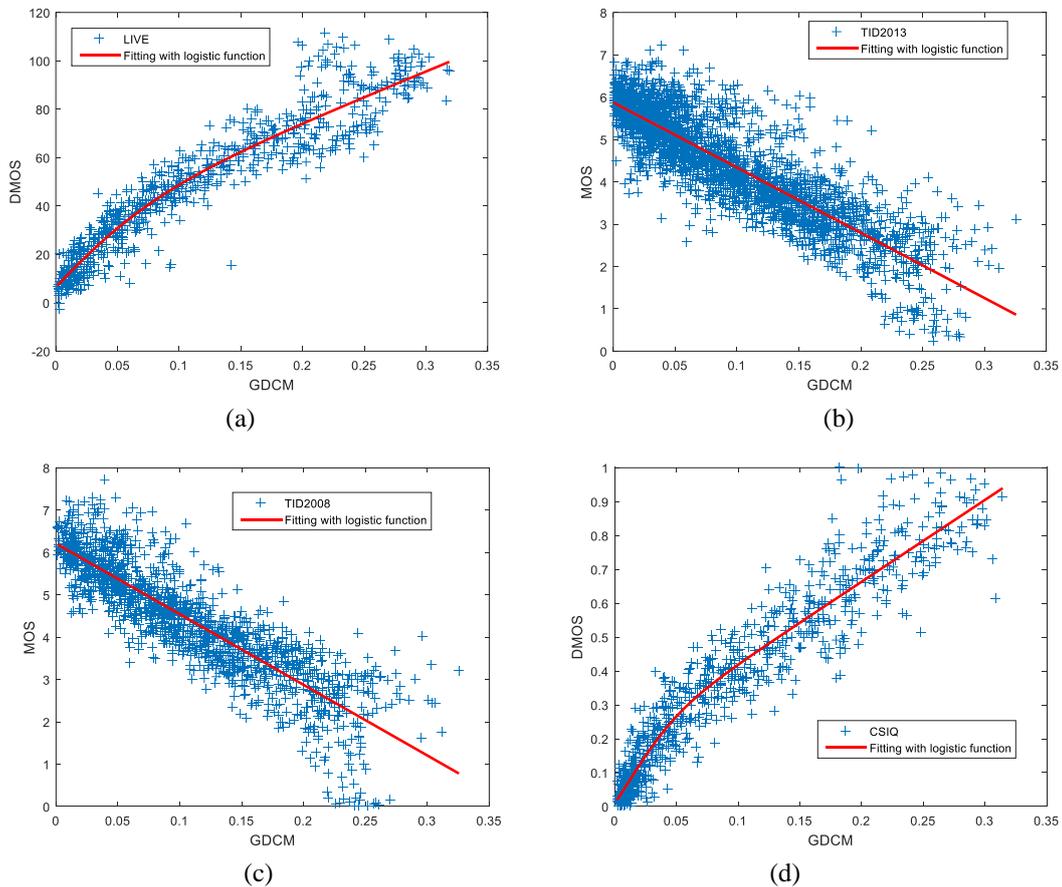

Figure 6: Scatter plots of subjective scores versus scores from the proposed scheme on IQA databases: (a) Live; (b) TID2013; (c) TID2008; (d) CSIQ.
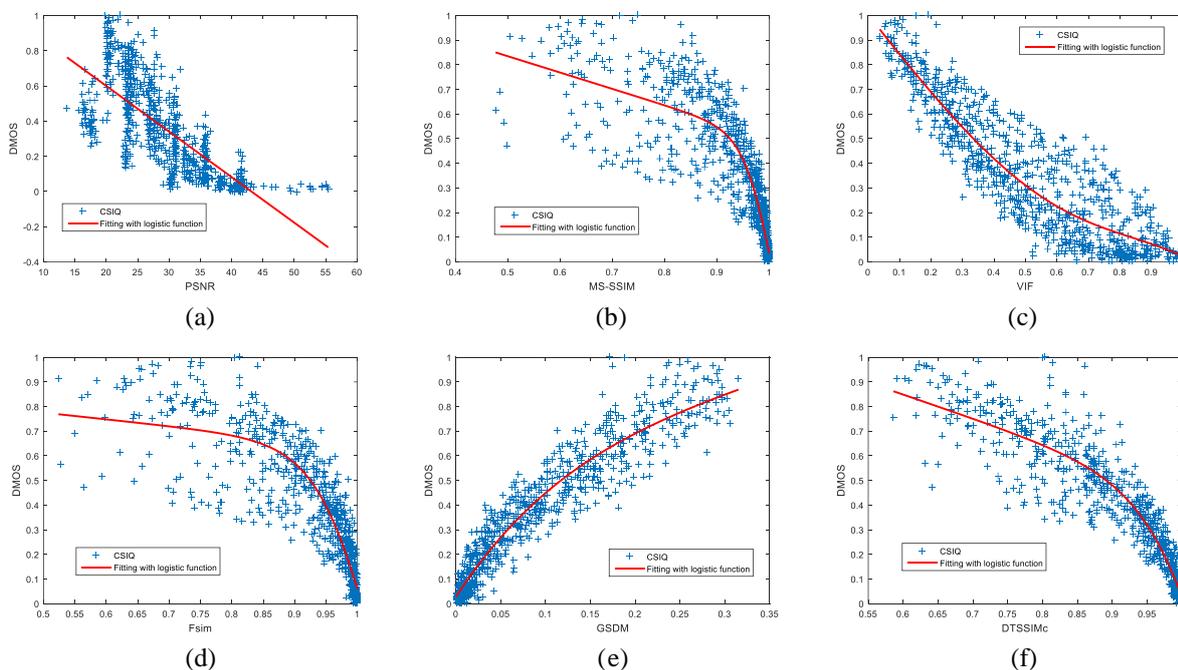

Figure 7: Scatter plots of subjective scores versus scores from some of the state-of-the-art algorithms on CSIQ database: (a) PSNR; (b) MS-SSIM; (c) VIF; (d) Fsim ;(e) GSDM (f) DTSSIMc.

CC=0.800/ROOC=0.801, Fsim values are CC=0.919/ROOC=0.931, GSDM values are CC=0.955/ROOC=0.957 and DTSSIMC values are CC=0.947/ROOC=0.956. PSNR values indicate that this later is not very well harmonized with perceived visual quality. The examination of the results obtained with MS-SSIM, lead us to say that GDCM has a significant performance on MS-SSIM.

It can be also observed from Table 8, that the proposed GDCM provides the best results on two out of four IQA benchmark databases. Moreover, it gives the second best CC and ROCC values on TID2013. Its three first places and especially the first ranking among well recognized methods in the literature confirms its location and its robustness and efficiency in the field of image quality assessment.

Figure 6 depicts the scatter plots of subjective scores

method. While, figure 7 shows the scatter plots of subjective scores versus scores from some of the state-of-the-art algorithms (VIF, MS-SSIM, PSNR, Fsim, GSDM and DTSSIMC) on CSIQ database. The red curves in Figure 6-7 are obtained by a nonlinear fitting based on Eq. (12). Here X-axis represents the objective score computed by the proposed metric and Y-axis represents the DMOS or MOS obtained through subjective tests. It is obvious that the sample points are much closer to the fitted curve in Figure 6, which indicates that objective scores predicted by GDCM, are more consistent with subjective assessment than other methods. Furthermore, this observation is confirmed by examining Figure 6 (d) and Figure 7: the sample points are much closer to the fitted curve in Figure 6 (d), Whereas, the sample points are less close to the fitted curve especially in figure 7 (a),

Table 3: Live database and performance comparison.

Method	RMSE	CC	KROCC	ROCC
IF	**7.667**	**0.960**	**0.827**	**0.963**
FC	10.264	0.927	0.758	0.926
IS-SSIM	9.259	0.941	0.804	0.951
SNR	10.469	0.924	0.763	0.928
SNR	13.360	0.872	0.687	0.876
SIM	8.945	0.945	0.796	0.948
IAD	**6.924**	**0.967**	**0.842**	**0.967**
sim	**7.530**	**0.961**	**0.836**	**0.965**
QM	11.193	0.912	0.741	0.909
CTex	8.990	0.944	0.807	0.948
SM	9.038	0.944	0.813	0.955
SDM	7.526	0.961	0.831	0.962
TSSIMC	*8.2754*	*0.9530*	*0.8182*	*0.9570*
VRO	**7.596**	**0.961**	**0.831**	**0.963**
DCM	8.227	0.954	0.822	0.959

Table 4: TID2008 database and performance comparison.

Method	RMSE	CC	KROCC	ROCC
VIF	0.789	0.809	0.587	0.750
IFC	0.911	0.734	0.424	0.568
MS-SSIM	0.717	0.845	0.657	0.854
VSNR	0.981	0.682	0.535	0.705
PSNR	1.100	0.573	0.421	0.579
SSIM	0.855	0.771	0.577	0.775
MAD	0.747	0.831	0.645	0.834
Fsim	**0.647**	**0.876**	**0.699**	**0.884**
NQM	1.059	0.614	0.461	0.624
DCTex	1.111	0.561	0.410	0.497
GSM	0.715	0.846	0.665	0.855
GSDM	**0.644**	**0.877**	**0.705**	**0.891**
DTSSIMC	**0.621**	**0.887**	**0.717**	**0.897**
GVRO	0.645	0.877	0.704	0.885
GDCM	**0.631**	**0.883**	**0.718**	**0.901**

Table 5: CSIQ database and performance comparison.

Method	RMSE	CC	KROCC	ROCC
VIF	0.098	0.928	0.754	0.920
IFC	0.143	0.838	0.590	0.767
MS-SSIM	0.115	0.899	0.739	0.913
VSNR	0.158	0.799	0.624	0.810
PSNR	0.158	0.800	0.598	0.801
SSIM	0.133	0.861	0.691	0.876
MAD	**0.082**	**0.950**	**0.797**	**0.947**
Fsim	0.103	0.919	0.769	0.931
NQM	0.176	0.743	0.564	0.740
DCTex	0.161	0.792	0.642	0.804
GSM	0.116	0.898	0.740	0.913
GSDM	**0.078**	**0.955**	**0.810**	**0.957**
DTSSIMC	**0.085**	**0.947**	**0.816**	**0.956**
GVRO	0.107	0.927	0.763	0.914
GDCM	**0.076**	**0.958**	**0.818**	**0.960**

Table 6: TID2013 database and performance comparison.

Method	RMSE	CC	KROCC	ROCC
VIF	1.032	0.772	0.515	0.677
IFC	0.902	0.554	0.394	0.539
MS-SSIM	0.839	0.833	0.605	0.786
VSNR	0.788	0.740	0.508	0.681
PSNR	0.761	0.669	0.470	0.640
SSIM	0.686	0.790	0.559	0.742
MAD	0.596	0.827	0.604	0.781
Fsim	**0.568**	**0.877**	**0.667**	**0.851**
NQM	0.943	0.686	0.474	0.643
DCTex	0.660	0.650	0.457	0.586
GSM	**0.698**	**0.846**	**0.626**	**0.795**
GSDM	**0.632**	**0.860**	**0.635**	**0.807**
DTSSIMC	**0.761**	**0.889**	**0.689**	**0.871**
GVRO	0.636	0.858	0.631	0.801
GDCM	**0.600**	**0.875**	**0.685**	**0.865**

against objective scores predicted by GDCM IQA

(b) and (d).

Table 7: Comparison of GDCM to the state-of-the-art. Direct average of measured CC and ROCC values are reported.

Method	*CC*	ROCC
VIF	0.867	0.828
IFC	0.763	0.700
MS-SSIM	0.880	0.876
VSNR	0.786	0.781
PSNR	0.729	0.724
SSIM	0.842	0.835
MAD	0.894	0.882
Fsim	0.908	0.908
NQM	0.739	0.729
DCTex	0.737	0.709
GSM	0.884	0.880
GSDM	**0.913**	**0.904**
DTSSIMC	**0.919**	**0.920**
GVRO	0.906	0.891
GDCM	**0.918**	**0.921**

Table 8: The position of all measures.

Measures	*TID2013*	*CSIQ*	*TID2008*	*LIVE*
VIF	11	6	10	3
IFC	15	14	14	13
MS-SSIM	7	8	7	9
VSNR	10	11	11	12
PSNR	13	13	13	15
SSIM	9	10	9	10
MAD	8	4	8	1
Fsim	**3**	5	5	2
NQM	12	15	12	14
DCTex	14	12	15	11
GSM	6	9	6	8
GSDM	4	**2**	**3**	5
DTSSIMC	**1**	**3**	**2**	7
GVRO	5	7	4	4
GDCM	**2**	**1**	**1**	6

5   Conclusion

In this work, the color distortion is introduced in *GSDM*. The color image coded in *RGB* color space in converted to another space including three components Y, I and Q. The deformed and gradient operator are computed from Y component, then the I and Q elements are used to compute the color distortion. The proposed measure is compared with twelve state-of-art methods. The performance of GDCM was the best compared with those measures. Some enhancements will be visited in the future works, including the investigation of others features.

References

[1] Z. Wang, E.P. Simoncelli, and A.C. Bovik, "Multi-scale structural similarity for image quality assessment," in Proc. IEEE Asilomar Conf. Signals, Syst., Comput., pp. 1398–1402. Pacific Grove, CA, 2003,
https://doi.org/10.1109/ACSSC.2003.1292216

[2] G.H. Chen, C.L. Yang, and S.L. Xie, "Gradient-Based Structural Similarity for Image Quality Assessment," in Proc. International Conference on Image Processing (ICIP06), pp. 2929-2932. Atlanta, GA, USA, 2006
https://doi.org/10.1109/ICIP.2006.313132

[3] Z.A. Seghir and F. Hachouf, "Edge-region information measure based on deformed and displaced pixel for Image Quality Assessment," Signal Processing: Image Communication 26 (8-9), pp. 534-549, 2011,
https://doi.org/10.1016/j.image.2011.06.003

[4] F. Zhang, L. Ma, S. Li, and K.N. Ngan, "Practical image quality metric applied to image coding," Transaction on Multimedia (IEEE) 13, pp. 615-624, 2011,
https://doi.org/10.1109/TMM.2011.2134079

[5] P.F. Felzenszwalb and D.P. Huttenlocher, "Distance Transforms of Sampled Functions," THEORY OF COMPUTING 8, pp. 415–428, 2012,
http://dx.doi.org/10.4086/toc.2012.v008a019

[6] P. Kovesi, "Image features from phase congruency," Videre: Journal of Computer Vision Research 1 (3), pp. 1–26, 1999,
https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.4.1641&rep=rep1&type=pdf

[7] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," Transactions on Image (IEEE) Processing 20 (8), pp. 1–26, 2011,
https://doi.org/10.1109/TIP.2011.2109730

[8] Z.A. Seghir and F. Hachouf, "Full-reference image quality assessment scheme based on deformed pixel and gradient similarity," Optik - International Journal for Light and Electron Optics 126 (24), 5946-5951, 2015,
https://doi.org/10.1016/j.ijleo.2015.08.132

[9] M. Gaubatz, "Metrix MUX Visual Quality Assessment Package: MSE, PSNR, SSIM, MSSIM, VSNR, VIF, VIFP, UQI, IFC, NQM, WSNR, SNR," [Online]. Available:
http://foulard.ece.cornell.edu/gaubatz/metrix_mux/

[10] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," IEEE Transactions on Image Processing 21(4), pp.1500–1512, 2012,
https://doi.org/10.1109/TIP.2011.2175935

[11] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," J. Electron. Imaging 19 (1), pp. 011006:1–011006:21, 2010, http://dx.doi.org/10.1117/1.3267105

[12] D. L.Ruderman, "The statistics of natural images," Netw. Comput. Neural Syst. 5 (4), pp. 517–548,1994, https://doi.org/10.1088/0954-898X_5_4_006

[13] C. Yang and S. H.Kwok, "Efficient gamut clipping for color image processing using LHS and YIQ. Journal of Optical Engineering 42 (3), pp. 701–711, 2003, http://dx.doi.org/10.1117/1.1544479

[14] H. R. Sheikh, Z. Wang, A. C. Bovik, and L. Cormack, "Image and Video Quality Assessment Research at LIVE 2005," [Online]. Available: http://live.ece.utexas.edu/research/quality.2005

[15] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008, A database for evaluation of full-reference visual quality assessment metrics," Advances of Modern Radioelectronics, Vol. 10, pp. 30-45, 2009, https://ponomarenko.info/papers/mre2009tid.pdf

[16] C. Larson and D. M. Chandler, "Categorical Image Quality (CSIQ) Database 2009," [Online]. Available: http://vision.okstate.edu/csiq . 2009

[17] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, and C.-C. Jay Kuo, "Color Image Database TID2013: Peculiarities and Preliminary Results," Proceedings of 4th Europian Workshop on Visual Information Processing EUVIP2013, Paris, France, June 10-12, pp. 106-111, 2013, https://ieeexplore.ieee.org/document/6623960

[18] VQEG report: www.its.bldrdoc.gov/vqeg/about-vqeg.aspx. 2010

[19] H.R. Sheikh, M.F. Sabir, and A.C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," Transaction on Image (IEEE) Processing 15 (11), pp. 3440-3451, 2006, https://doi.org/10.1109/TIP.2006.881959

[20] A. Zeggari, Z. A. Seghir, and M. Hemam, "Perceptual image quality assessment based on gradient similarity and Ruderman operator," Int. J. Comput. Vis. Robotics 11(2): 151-174, 2021, https://dx.doi.org/10.1504/IJCVR.2021.113402

[21] Z.A. Seghir, F. Hachouf, M. Hemam, F. Morain-Nicolier, and A. Zeggari, "Quality Evaluation Algorithm: New Structural Similarity by Using Distance Transform Approach and Gradient Similarity," in 2018 International Conference on Signal, Image, Vision and their Applications (SIVA). IEEE, 2018, https://doi.org/10.1109/SIVA.2018.8661089

[22] H.R. Sheikh, A.C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," IEEE Trans. Image Process., vol. 14, no. 12, pp. 2117-2128, Dec, 2005, https://doi.org/10.1109/TIP.2005.859389

[23] H.R. Sheikh and A.C. Bovik, "Image information and visual quality," IEEE Trans. Image Process., vol. 15, no. 2, pp. 430-444, Feb,2006, https://doi.org/10.1109/TIP.2005.859378

[24] D.M. Chandler and S.S. Hemami, "VSNR: a wavelet-based visual signal-to-noise ratio for natural images," IEEE Trans. Image Process., vol. 16, no. 9, pp. 2284-2298, Sep, 2007, https://doi.org/10.1109/TIP.2007.901820

[25] Z. Wang, A. C. Bovik, and L. Lu, "Why is image quality assessment so difficult?," In Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing, Orlando, May, 2002, http://dx.doi.org/10.1109/ICASSP.2002.1004620

[26] N. Damera-Venkata, T.D. Kite, W.S. Geisler, B.L. Evans, and A.C. Bovik, "Image quality assessment based on a degradation model," IEEE Trans. Image Process., vol. 9, no. 4, pp. 636-650, Apr, 2000, https://doi.org/10.1109/83.841940

[27] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600-612, Apr, 2004, https://doi.org/10.1109/TIP.2003.819861

[28] Y. Horita, K. Shibata, and Z. M. Parvez Saddad, "Subjective quality assessment toyama database," [Online]. Available: http://mict.eng.u-toyama.ac.jp/mict/ .2008

[29] A. Ninassi, P. Le Callet, and F. Autrusseau, "Subjective Quality Assessment-IVC Database 2005," [Online]. Available: http://www2.irccyn.ecnantes.fr/ivcdb. 2005

[30] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad, "MICT Image Quality Evaluation Database 2000," [Online]. Available: http://mict.eng.u-toyama.ac.jp/mict/index2.html. 2000

[31] D. M. Chandler and S. S. Hemami, "A57 Database 2007," [Online]. Available: http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.htm. 2007

[32] U. Engelke, H.-J. Zepernick, and M. Kusuma, "Wireless Imaging Quality Database," [Online]. Available: http://www.bth.se/tek/rcg.nsf/pages/wiq-db. 2010

[33] Z.A. Seghir and F. Hachouf, "Image Quality Assessment based on Edge-region information and Distorted pixel for JPEG and JPEG2000," in Proc. ACIVS 2009, Bordeaux, France, Sept.28 - Oct.2, LNCS 5807, pp. 156-166, 2009, https://doi.org/10.1007/978-3-642-04697-1_15

[34] Z.A. Seghir and F. Hachouf, "Edge-region information with distorted and displaced pixels measure for image quality evaluation," in Proc. ISSPA 2010, Kuala Lumpur, Malaysia, May 10- 13, pp. 77-80, 2010, https://doi.org/10.1109/ISSPA.2010.5605504

[35] Z.A. Seghir, F. Hachouf, and F. Nicolier, "Distance transform measure based on edge region information: An algorithm for image quality assessment," in Proc. NaBIC2011, Salamanca, Spain, Oct 19-21, pp. 125-130, 2011, https://doi.org/10.1109/NaBIC.2011.6089447

[36] Z.A. Seghir and F. Hachouf, "Image quality assessment scheme based on gradient similarity and color distortion," 12th International Symposium on Programming and Systems (ISPS) , Pages: 1 - 8, 2015, https://doi.org/10.1109/ISPS.2015.7244985

[37] Z.A. Seghir and F. Hachouf, "Color Image Quality Assessment Based on Gradient Similarity and Distance Transform," ACIVS 2015: 591-603, 2015, https://doi.org/10.1007/978-3-319-25903-1_51

[38] Z.A. Seghir and F. Hachouf, "Full-Reference Image Quality Assessment Measure Based on Color Distortion," in Proc. 5th IFIP TC 5 International Conference, CIIA Saida, Algeria, May 20–21, 2015, pp. 66–77, 2015, https://doi.org/10.1007/978-3-319-19578-0_6

[39] D.B. Eddine, F. Hachouf, and Z.A. Seghir, , "A New No-Reference Image Quality Assessment Based On SVR Fusion," in Proc. 5th European Workshop on Visual Information Processing (EUVIP), France, Dec 10 – 12, pp. 1-6, 2014, https://doi.org/10.1109/EUVIP.2014.7018390

[40] Z.A. Seghir, F. Hachouf, and F. Nicolier, "Blind Image Quality Metric for Blurry and Noisy Image," in Proc. IEEE Second International Conference on Image Information Processing, Jaypee University of Information Technology, India, Dec 9–11, pp. 193-197, 2013, https://doi.org/10.1109/ICIIP.2013.6707581

[41] N. Thakur, N. Uddin Khan, and S. D. Sharma, "A Review on Performance Analysis of PDE Based Anisotropic Diffusion Approaches for Image Enhancement," Informatica 45, pp. 89–102, 2021, https://doi.org/10.31449/inf.v44i6.3333

[42] N.H.Hai, L.M.Hieu, D.N.H. Thanh, N. V. Son, P. V.B. Surya, "An Adaptive Image Inpainting Method Based on the Weighted Mean," Informatica 43 ,pp 507–513, 2019, https://doi.org/10.31449/inf.v43i4.2461

[43] D.Thanh, P.Surya, L.Hieu, "A Review on CT and X-Ray Images Denoising Methods," Informatica 43, pp 151–1592019, https://doi.org/10.31449/inf.v43i2.2179

[44] T.Zhang, D. Li, Y. Cai, Y. Xu, " Super-resolution Reconstruction of Noisy Video Image Based on Sparse Representation Algorithm," Informatica 43, pp 415–420, 2019, https://doi.org/10.31449/inf.v43i3.2916