Abstract—Objective image quality assessment (QA) is a fundamental and challenging job in image processing which evaluates the image quality consistently with human perception automatically. Generally, an image can be segmented into two kinds of areas: structure and texture. And structural information plays a much more important role between the two. The pixels, edges and shape with directional characteristic contribute much more to the structural information. In this paper, the structural information is modeled as the energy of directional projection, which is shown in a directional projection-based map that is built by Radon transform. On the assumption that any image’s distortion could be modeled as the difference between the directional projection-based maps of reference and distortion images, we propose a new objective quality assessment method with Radon transform for full reference model, and we also try to explore the feasibility to develop the image QA metrics with scalable efficiency and computational cost. Experimental results show that the proposed Metrics are well consistent with the subjective quality score.

Keywords—image quality assessment, signal projection, Radon transform, objective quality

I. INTRODUCTION

In the procedure of image and video processing system, e.g., acquisition, processing, coding, storage, transmission and reproduction, digital image and video may be in de-gradation in visual quality. The purpose of the research of image quality assessment (QA) is to develop strategies and algorithms to evaluate the quality accurately consistently with subjective perception, which is a challenging and fundamental job. Image QA is also with many interests in many applications such as: dynamic monitoring and adjusting image quality, optimizing algorithms and parameter settings of image processing systems, and benchmarking image processing systems and algorithms [1]. In many cases, quality measure methods with scalable efficiency and computational cost are desirable, especially for some real-time or high-performance applications, e.g., image retrieval [2], [3], which is worth investigating.

Image QA can be classified as subjective and objective image QA [4]. The subjective QA is more accurate. However, in practice, it is usually expensive, time-consuming, inconvenient, and environment-limited. Moreover, this kind of method may be affected by various of factors, for example, the mood of the candidate, the status of testing equipment, the individuality of observers, and so on. Therefore, it is important to develop an objective image quality measure which automatically and exactly evaluates the image quality. In this paper, we focus on full-reference image QA, which means that the original image (we should take it grand that the original image is ‘perfect’ or of ‘high quality’) is completely known as the reference one.

In the past three decades, a lot of objective image quality assessment methods have been put forward [1], [4], [5], which can be generally classified into the following three categories:

- Metrics based on the statistic of pixels, which use some mathematical statistic to represent the quality of the image, such as the mean square error (MSE), the root mean square error (RMSE), the signal to noise ratio (SNR), and the peak signal to noise ratio (PSNR). However, these mathematically defined metrics above cannot completely agree with human’s perception [1], [6], [7]. Although these methods have some systematic drawbacks, they are still widely used in many situations, since they are independent to the images and easy to calculate. The QA research aims to improve on the PSNR.

- Metrics based on human visual systems (HVS). The idea based on the characteristics of human vision to the image QA is put forward by Mannos and Sakrison in a famous paper [8] in 1974. Some other people also contribute a lot in this field [8]-[14]. Although the image QA metric based on the psychophysical measurement of HVS is mostly accepted, the complexity of the HVS and the finitude of the cognizing of the human beings still keeps this metric from going much further [1], [15].

- Metrics based on structural distortion of images. This idea is brought forward by Wang et al. [1]. On the assumption that “human visual perception is highly adaptive for extracting structural information from a scene”, they propose the “Structural Similarity (SSIM)” metric which compares the structural similarity between the original image and the distorted one. The arithmetic of Mean Structural Similarity

Authorized licensed use limited to: IEEE Xplore. Downloaded on May 6, 2009 at 04:50 from IEEE Xplore. Restrictions apply.
(MSSIM) is that: the image is divided into small blocks, and the distortion of each block is computed using the information of luminance, contrast and structure, and the final quality score of the image is the mean value of the blocks [1]. Shnayderman et al. propose an idea of assessing the image based on the singular value decomposition (SVD) of the matrix of the image which gets a good result [16]. These metrics above are of simple parameters and low computational complexity, which are potential to replace the role of those mathematically defined metrics. Some new measures are brought out recently [17]-[20].

Generally, the developed image quality measure evaluates the quality of images to achieve the agreement with human perception, also it would be adapted to both individual and cross distortion types, i.e., it is universal and does not depend on testing images, testing environment and the observers individually. Moreover, the image quality measure with high efficiency and low computational complexity are also desired for some image and video processing applications. Although mathematically defined metrics have those properties, they do not have a satisfactory performance to be consistent with human perception. In our previous work [21], a projection-based image quality metric with simple parameters is proposed. This work is to develop a universal measure with scalable efficiency and computational complexity using simple parameters, and our approaches would work as an expansion for those mathematically defined ones.

The paper is organized as follows: Section II introduces the proposed method; In Section III, our experimental results are compared with some other metrics. Section IV concludes.

II. THE PROPOSED METRICS

In a general way, an image may be segmented into two kinds of areas: structure and texture, and especially between them, structural information plays a more important role usually. Natural images are of much structure (contains much important information about the objects' structure for the visual scene [1]). The pixels, edges and shape with directional characteristic contribute a lot to the structural information. The distortion of the structural information could be modeled as the alteration of the directional characteristic.

When the image is degraded, the directional characteristic could be distorted. In this paper, we try to use a directional projection-based map to present the directional characteristic, and the map is built by Radon transform [22] with different parameters. The maps with different parameters perform the scalable directional characteristic and the scalable efficiency and computational cost.

Firstly the Radon transform is introduced simply. The Radon transform [22] can be defined by:

$$\mathcal{R}(s, \theta)[f(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(s - x \cos \theta - y \sin \theta) dx dy$$

(1)

where \(f(x, y)\) is a 2-D vector, \(s\) is the perpendicular distance from a line to the origin and \(\theta\) is the angle formed by the distance vector. Radon transform could be regarded as the projection procedure on different directions, so we name this kind of method the “directional projection”. \(\mathcal{R}(s, \theta)\) is the function of Radon transform and \(\mathcal{R}(s, \theta)[f(x, y)]\) is the vector of the directional projection-based map.

The reference image is divided into small blocks with the block size \(m \times l\), and the \(n\)th block is defined as the vector \(\mathbf{b}_n \in \mathbb{R}^{m \times l}\), and \(\mathbf{d}_n \in \mathbb{R}^{m \times l}\) is defined as the vector of the directional projection-based map. \(\mathbf{b}_n \in \mathbb{R}^{m \times l}\) and \(\mathbf{d}_n \in \mathbb{R}^{m \times l}\) is defined for the counterparts of the distortion image. \(\mathbf{D}_n\) and \(\mathbf{d}_n\) is calculated as follows:

$$\mathbf{D}_n = \mathcal{R}(s, \theta)[\mathbf{b}_n]$$

(2)

$$\mathbf{d}_n = \mathcal{R}(s, \theta)[\mathbf{b}_n]$$

(3)

Define the local distortion intensity as \(SD_n\):

$$SD_n = \| \mathbf{b}_n - \mathbf{d}_n \|$$

(4)

where \(\| \cdot \|\) represents the procedure of calculating the vector norm.

The global distortion intensity \(SD\) is simply calculated as the mean of the local distortion intensity.

$$SD = \text{mean}(SD_n)$$

(5)

Here we are cautious to propose that our predictive score of objective quality is a logarithmic function of the distortion intensity which obeys the Weber-Fechner law [23] (a constant relative difference in the intensity corresponds to a constant absolute difference in the logarithm of the intensity). Therefore, the image quality measure with directional projection (DP) is defined:

$$DP = \log(SD)$$

(6)

and it is clear that \(SD>0\).

In this work we build the directional projection-based maps based on Radon transform. The differences of the maps are desired to represent the variation of the images’ degradation, which model any image distortion. The actual value is meaningless, but the comparison between two values for different distorted images gives one measure of quality. The lower the predicted score of \(DP\) is, the better the image quality is. When the distortion and the reference images are identical, \(SD = 0\).

In the Eq.(1), there are two parameters \(s\) and \(\theta\), and \(\theta\) could be used as the scale of the efficiency and computational cost. \(\theta\) is varied to adapt to the different applications, and achieve the scalable efficiency and computational cost.
Figure 1. The flow chart of DP

Figure 2. The directional projection-based map of one image.

Figure 1 shows the flow chart of the algorithm. Figure 2 gives a directional projection-based map of one image as an example.

EXPERIMENTAL RESULTS AND DISCUSSION

The database we use in our experiments is the known "LIVE Image Quality Assessment Database Release 2" [24], and the database is composed of color nature images. The subjective score of the images (DMOS, Difference Mean Opinion Score) comes from the latest database [14]. Some images in the database are randomly selected in Figure 3 as an example.

The database includes 29 reference color images, each of which contains 5 distortion types (total 799 images): Fast Fading Rayleigh (FF, 145 images), Gaussian Blur (GBlur, 145 images), White Noise (WN, 145 images), JPEG (175 images), and JPEG2000 (169 images). The five distortion types, which could often happen in practical applications, are introduced into studies in this work. FF is a simulation of transmission errors in compressed JPEG2000 bit stream using a Fast fading Rayleigh channel model. The RGB components are blurred using a circular-symmetric 2-D Gaussian kernel in GBlur distortion. WN distorts the images by adding white Gaussian noise to RGB components. JPEG and JPEG2000 compress the images at different bit rates, which could often happen in image and video processing applications. We evaluate the performances following the procedures in the Video Quality Experts Group (VQEG) Phase I FR-TV test [25]. And the simple and widely used metric, namely PSNR, and other two, namely MSSIM [1] and SVD [16] are chosen to compare with our metrics.

Figure 3. Some example images from the database (all images are resized and converted into grayscale image for visibility)

Figure 4: Scatter plots for PSNR, SVD, MSSIM, and DP for the distortion of 5 types

Figure 5: Scatter plots for PSNR, SVD, MSSIM, and DP for JPEG and JPEG2000

In our experiment, we choose $m \times l = 8 \times 8$, just because it is a common size in many image processing applications and both SVD and MSSIM use this window size. The experiments...
work with the luminance of the images. We convert color images into grayscale ones by separating the luminance information from the color information.

Particularly, we compare three DP with different parameter $\theta$. DP1 and DP2 are desired to achieve scalable efficiency and computational cost.

$$\begin{align*}
\text{DP: } & \theta = 0 : \pi - 1 \\
\text{DP1: } & \theta = 45 : \pi - 1 \\
\text{DP2: } & \theta = 30 : \pi - 1
\end{align*}$$

A. Experimental Results

In Figure 4 and Figure 5, the X-axis is the predictive score of each assessment metric and the Y-axis is DMOS. The lines in figures are non-linear fitting curves which are chosen for regression or fitting for each of those methods, and the logistic function is with five variables as follows:

$$\text{logistic}(x) = \frac{a_0 + a_4 - a_3}{1 + \exp(\frac{x - a_2}{a_1})} \tag{7}$$

Figure 4 is the DP results for all images comparing the performances between each of the four metrics and DMOS in cross-type distortions. Figure 5 is the DP results for JPEG and JPEG2000 which compare the performances of cross image coding types. Table I and Table II compare the Spearman rank correlation-coefficient (SROCC) and the Pearson correlation-coefficient (PCC) between each of the four metrics and DMOS. Table 3 compares RMSE between each of the four metrics and DMOS.

B. Discussion

From figures and tables above we come to the conclusion that PSNR is not well-adaptable in all of the distortion types except WN. Meanwhile, it is reasonable that PSNR has a satisfactory performance of all in WN, since the information of the WN-distortion image is distorted only by WN, and thus PSNR can count more accurately these errors which are statistically independent. When the “errors” or characters which distort the images are not uncorrelated, PSNR cannot work well simply and accurately. In this case, the other metrics try to overcome some systematic drawbacks of PSNR.

SVD and MSSIM have close performances, while MSSIM shows the best of all in FF. In individual distortion type, DP outperforms the others in GBlur, JPEG and JPEG2000, and it also has a good performance in FF and WN. DP has the best performance in cross-distortion types, especially in coding types.

There are some issues which are worth investigating. All of the four metrics only work with the luminance of the images. However subjective score DMOS are gained by the observers evaluating the color images. Therefore, when the distortion of color information, which may not be detected in luminance channel, happens, it is much difficult to assess those images exactly only by using luminance information. In this database, FF is the distortion which sometimes degrades the color information, so the plots of four metrics scatter in FF. It will not be an easy job to study color distortion for image quality. Moreover, the sensitivity of DP to slight distortions in rotation, shift and magnification is not satisfactory, which is to be taken into future research.

DP has a more computational complexity compared with SVD and MSSIM, but the computational complexity of DP1 and DP2 with different parameters has been significantly improved. The implementation of DP1 and DP2 on a 768 $\times$ 512 image (bikes.bmp) on a Pentium IV, 3.0GHz laptop using the luminance information takes about 0.2 second. Besides, the typical DP values range between -6 and 4 in our implementations.

III. CONCLUSION AND FUTURE WORK

On the assumption that any image distortion can be modeled as the difference of the directional projection-based maps, we propose an objective image measure based on directional projection by using Radon transform. The experimental results show that DP outperforms PSNR, SVD, and MSSIM. This metric is not only well adaptable in individual distortion type, especially in image coding types, but also in cross-distortion types.

Our future work is to explore into more aspects about the relationship between the image distortion and the subjective visual quality, and we will also focus on the research of extending the proposed metric to color images and video sequences.

REFERENCES


**TABLE I.** CC-BASED FOR PSNR, SVD, MSSIM AND PD WITHIN INDIVIDUAL DISTORTION AND CROSS TYPES

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>FF</th>
<th>GBlur</th>
<th>WN</th>
<th>JPEG</th>
<th>JPEG2000</th>
<th>JPEG+JPEG2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>0.8693</td>
<td>0.8936</td>
<td>0.7734</td>
<td>0.9844</td>
<td>0.8865</td>
<td>0.8980</td>
<td>0.8863</td>
</tr>
<tr>
<td>SVD</td>
<td>0.8822</td>
<td>0.8985</td>
<td>0.7220</td>
<td>0.9786</td>
<td>0.9589</td>
<td>0.9428</td>
<td>0.9466</td>
</tr>
<tr>
<td>MSSIM</td>
<td>0.8984</td>
<td>0.9422</td>
<td>0.8465</td>
<td>0.9699</td>
<td>0.9482</td>
<td>0.9407</td>
<td>0.9377</td>
</tr>
<tr>
<td>PD</td>
<td>0.9285</td>
<td>0.9029</td>
<td>0.9132</td>
<td>0.9003</td>
<td>0.9734</td>
<td>0.9467</td>
<td>0.9585</td>
</tr>
<tr>
<td>PD1</td>
<td>0.9275</td>
<td>0.9036</td>
<td>0.9097</td>
<td>0.9803</td>
<td>0.9731</td>
<td>0.9462</td>
<td>0.9581</td>
</tr>
<tr>
<td>PD2</td>
<td>0.9283</td>
<td>0.9033</td>
<td>0.9119</td>
<td>0.9804</td>
<td>0.9731</td>
<td>0.9465</td>
<td>0.9583</td>
</tr>
</tbody>
</table>

**TABLE II.** SROCC-BASED FOR PSNR, SVD, MSSIM AND PD WITHIN INDIVIDUAL DISTORTION AND CROSS TYPES

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>FF</th>
<th>GBlur</th>
<th>WN</th>
<th>JPEG</th>
<th>JPEG2000</th>
<th>JPEG+JPEG2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>0.8744</td>
<td>0.8939</td>
<td>0.7709</td>
<td>0.9831</td>
<td>0.8798</td>
<td>0.8931</td>
<td>0.8890</td>
</tr>
<tr>
<td>SVD</td>
<td>0.8872</td>
<td>0.8989</td>
<td>0.7055</td>
<td>0.9835</td>
<td>0.9492</td>
<td>0.9404</td>
<td>0.9474</td>
</tr>
<tr>
<td>MSSIM</td>
<td>0.9075</td>
<td>0.9394</td>
<td>0.8595</td>
<td>0.9645</td>
<td>0.9432</td>
<td>0.9357</td>
<td>0.9403</td>
</tr>
<tr>
<td>PD</td>
<td>0.9312</td>
<td>0.9018</td>
<td>0.9070</td>
<td>0.9770</td>
<td>0.9602</td>
<td>0.9416</td>
<td>0.9575</td>
</tr>
<tr>
<td>PD1</td>
<td>0.9300</td>
<td>0.9014</td>
<td>0.9024</td>
<td>0.9767</td>
<td>0.9600</td>
<td>0.9416</td>
<td>0.9573</td>
</tr>
<tr>
<td>PD2</td>
<td>0.9309</td>
<td>0.9013</td>
<td>0.9052</td>
<td>0.9770</td>
<td>0.9603</td>
<td>0.9418</td>
<td>0.9576</td>
</tr>
</tbody>
</table>

**TABLE III.** RMSE-BASED FOR PSNR, SVD, MSSIM AND PD WITHIN INDIVIDUAL DISTORTION AND CROSS TYPES

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>FF</th>
<th>GBlur</th>
<th>WN</th>
<th>JPEG</th>
<th>JPEG2000</th>
<th>JPEG+JPEG2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>10.1909</td>
<td>12.2479</td>
<td>7.5269</td>
<td>5.5200</td>
<td>7.2996</td>
<td>8.1296</td>
<td>8.4142</td>
</tr>
<tr>
<td>PD1</td>
<td>10.2789</td>
<td>12.2046</td>
<td>7.6691</td>
<td>5.5200</td>
<td>7.3427</td>
<td>8.1628</td>
<td>8.4495</td>
</tr>
<tr>
<td>PD2</td>
<td>10.2109</td>
<td>12.2211</td>
<td>7.5810</td>
<td>5.5170</td>
<td>7.3366</td>
<td>8.1391</td>
<td>8.4317</td>
</tr>
</tbody>
</table>