Directional Regularity for Visual Quality Estimation

Delei Liu, Yong Xu*, Yuhui Quan, Zhiwen Yu

School of Computer Science & Engineering, South China University of Technology, Guangzhou 510006, China

Patrick Le Callet

the EURASIP membership, IRCCyN Lab, University of Nantes, BP 50609 44306 Nantes Cedex 3, France

Abstract

In this paper, a reduced-reference image quality assessment metric is proposed, which measures the difference of the regularity of the projection spatial arrangements between the reference image and the distorted image. The projections of an image are first extracted by Radon transform. Then fractal dimensions are computed on each projection and concatenated as the image features that characterize the image structures from the view of the spatial distribution. Finally the image features are pooled as the quality score using ℓ_1 distance. The proposed approach was evaluated on four public benchmark databases. Experimental results have demonstrated the excellent performance of the proposed approach.

Keywords: reduced-reference image quality assessment, fractal dimension, Radon transform

*Corresponding author: Yong Xu

Email addresses: liu.dl@mail.scut.edu.cn (Delei Liu), yxu@scut.edu.cn (Yong Xu), yuhui.quan@mail.scut.edu.cn (Yuhui Quan), zhwyu@scut.edu.cn (Zhiwen Yu), patrick.lecallet@univ-nantes.fr (Patrick Le Callet)

Project supported by National Nature Science Foundations of China (61273255, 61211130308 and 61070091), Fundamental Research Funds for the Central Universities (SCUT 2013ZG0011) and GuangDong Technological innovation project (2013KJCX0010).

1. Introduction

Action recognition is a widely studied topic in computer vision. It has benefited a wide range of applications such as video surveillance, human-computer interaction and video retrieval. In applications, action recognition may be affected by the image degradation, such as image blur, compression and transmission. In other words, action recognition from images can benefit from high image quality. How to quantify image quality in visual meaning for human visual system (HVS), leading to the research on image quality assessment (IQA). IQA can be categorized into two groups: subjective IQA and objective IQA. Subjective IQA is simple and accurate, however, it can not work without human's participation. Furthermore, it is infeasible in the case that there are a mass of images. Therefore, it is indispensable to attach more importance to objective IQA.

According to the availability of a reference image, objective IQA can be further divided into three categories: full-reference (FR), no-reference (NR) and reduced-reference (RR). FR-IQA [1–5]considers that the reference image is perfect and is available to access. A test image quality is assessed by calculating the similarity between the test image and its reference image. However, the reference image is always not available. NR-IQA [6–8]is not relative to the reference image. From the view of applications, NR metrics are more adaptable to different conditions. But in fact, due to the varied image contents and the individual distortion types, NR-IQA is extremely difficult when no prior knowledge is available. As a compromise between FR-IQA and NR-IQA, RR-IQA [9–13]is designed, which uses partial information of the reference image for quality prediction. This paper focuses on RR-IQA.

RR-IQA expects to use less data of the reference image and achieve higher prediction accuracy. Thus, on the way to RR-IQA metric, the challenging task is the selection of the smallest set of features that can support quality assessment effectively. Therefore, studying and exploiting the special properties of natural

images has been one of the most important tasks in RR-IQA. The recent focus of RR-IQA research mainly depends on some statistic models of natural images (*e.g.*, [9, 11, 13–17]). However, these RR-IQA metrics either need a large amount of reference data to achieve good performance, or have high computational complexity. To this end, in this paper we develop an efficient RR-IQA model based on Radon-representation and fractal analysis, called Similar Regularity of Radon-representation Measure(SRRM). Our approach not only can efficiently encode both structural information and spatial information, which is much more important for IQA, but also has low computational complexity.

Recent years, despite lots of work on RR-IQA, some problems remain unsolved. The first is the structural information of images. As we known, the structural information can have a dramatic impact on IQA [18], but many existing methods cannot handle this problem. For instance, Singular Value Decomposition(SVD) [19], the singular values have been used for image features set, which loses the image's structural information [20]. The second is the spatial information of images. Different part of an image has different contribution to the image quality. Many researchers does not consider this information. Such as the statistic models of natural images [9, 11, 13–17] that lose the spatial information. The third is the high data rate. A successful RR-IQA is expected to use less data of images and achieve higher evaluation accuracy. Many RR-IQA metrics need a large amount of image's information to achieve good performance. For instance, in [12], features are extracted in the size of 3×3 windows and finally concatenated as image feature set. The fourth is the high computational complexity. From the view of applications, a RR-IQA model should be not only effective but also efficient. However, some RR-IQA metrics have high computational complexity. For example, in [11], a number of parameters that need to be trained on databases.

In this paper, we propose the novel RR-IQA metric, SRRM that can handle above problems. First, SRRM converts original pixel represented images into

Radon-pixel images by using the Radon transform [21]. The new Radon representation is more structural informative in geometry and has a much lower dimension. Second, SRRM characterizes Radon-pixel by fractal dimension(FD) which not only has a strong correlation with HVS [22], but also can encode the image's spatial information[23–27]. Third, fractal analysis is run on 1D projected signal of the image, which has low computational complexity.

The proposed RR-IQA feature characterizes the spatial regularity of the distribution for the Radon-pixel image based on fractal analysis. The Radon representation is first extracted by Radon transform. FD is then used to measure the irregularity of the spatial distribution of the image's projection. Finally all the computed FDs are concatenated as a feature vector. By using fractal analysis that has a strong correlation with HVS [22], the image structures are well represented and the irregularity of spatial difference between reference image and distorted image can be well characterized. Our approach was evaluated on four famous benchmark IQA databases using five popular evaluation criteria. The competitive results achieved demonstrate that our method performs on par with the state-of-the-art approaches.

SRRM has several advantages as follows.

- It is able to handle the structural information of images;
- It is able to handle the spatial information of images;
- It is easy to use. There is only one parameter in SRRM and does not need careful adjustment.
- It is effective. It yields satisfactory results in our experiments.

The rest of this paper is organized as follows: Section 2 is devoted to related work. Section 3 gives a brief review on the Radon transform and fractal analysis. A detailed description of our proposed metric is given in Section 4. Experimental

evaluation and result analysis are presented in Section 5. Finally, Section 6 concludes the paper.

2. Related Work

In the past decade, many RR-IQA methods have been put forward, we give a brief review in this section.

In previous work, RR-IQA focuses on the specific types of image distortions, these methods use image distortion modeling [17, 28] developed for specific application environments. For instance, in [28], a hybrid image quality metric combines five structural features: blocking, blur, edge-based image activity, gradient-based image activity and intensity masking. The metric makes a similarity assessment between the distorted image and the reference image to assess JPEG coded images. However, these metrics suffer from bad performance when images with different distortion types are tested together, because the models are built for each distortion type respectively.

Recent work has concentrated on general-purpose RR-IQA methods. Most of these approaches are base on natural image statistical model [11, 13, 14, 16] and have achieved impressive results in RR-IQA. The basic idea of these methods is to quantify the image quality by quantifying the disturbance to the image statistics caused by the distortion. Wang *et al.*[13] modeled natural images using the marginal probability distributions of the coefficients in wavelet domain, and the Kullback-Leibler distance (KLD) between two marginal distributions is used to measure the image distortion. In order to model the perceptual sensitivity of biological vision, Li *et al.*[11] proposed the so-called divisive normalization transformation (DNT) for image representation. The image statistic is based on the Gaussian scale mixtures(GSM) model and the KLD is used to pool the features to the final score. In [9], according to the distribution of wavelet coefficients, geometric information is extracted for quality assessment. In [15], the

generalized Gaussian density is employed to model the distribution of the discrete cosine transform (DCT) coefficients. Xue *et al.*[14] employed the Weibull distribution to describe the statistics of image gradient magnitude. In [12], the image quality is measured by the difference between the entropies of wavelet coefficients of reference and distorted images. To adapt the SSIM[18] to RR-IQA, Rehman *et al.*[1] combined the GSM-based statistics in a multi-scale and multi-orientation DNT domain following the philosophy in the construction of SSIM. A regression-by-discretization method is then applied to normalize the measure across image distortion types. All of these methods are based on counting the difference of the numbers of elements in two images, which lose the details of how the elements are distributed.

Another most interesting direction of RR-IQA is the usage of the Singular Value Decomposition(SVD) [19]. The singular values calculated for 8×8 pixels blocks of reference and distorted images have been used for the calculation of the quality factors of each block which are proportional to the square roots of the aggregated squared differences of 8 singular values. The overall image quality score has been defined as the mean difference between the quality factors computed for each block and the middle element of the sorted vector of factors. The drawback of this metric is that it loses the image structure information [20].

In this paper, we develop an efficient RR-IQA model based on Radon-representation and fractal analysis, our approach not only can efficiently encode the structural information, but also characterize the spatial information.

3. Preliminaries

Before presenting the detailed description of our approach, we give an introduction of two mathematical tools upon which our approach is built. we first describe the Radon transform. Then we introduce fractal analysis which encodes coefficients for Radon transform, which can offer us the ability to precisely evaluate the visual structural information of images.

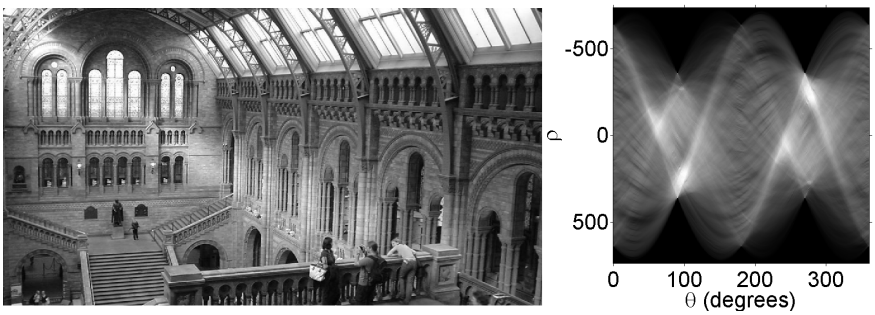

Figure 1: The image and its Radon transform coefficients. Left: an image"Railway Station" in LIVEmd database. Right: its Radon transform of the image.

3.1. Radon Transform

The Radon transform [21] computes projections of an image matrix along specified directions. It is defined as follows:

$$R_\theta(\rho) = \int_{-\infty}^{+\infty} f(x'cos\theta - y'sin\theta, x'sin\theta + y'cos\theta)dy', \tag{1}$$

where

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The Radon operator maps the spatial domain $f(x, y)$ to the projection domain (ρ, θ), in which θ is the angle and ρ is the signed distance to the origin of the coordinate system. The Radon transform for a large number of angles is often displayed as an image. Figure 1 shows the example of Radon transform of the image "Railway Station" which is taken from LIVEmd database. We can see that some very bright spots are found in the Radon transform plot, the location of each bright spot corresponds to one or many strong lines in the original image.

In this work, we have two major reasons to adopt the Radon transform. First, the Radon-pixel image brings a large advantage to achieving global geometric structural information. This is because the Radon-pixel image has more

geometric information than the original pixel image. It can be seen that one Radon-pixel corresponds to a line segment which needs two pixels in the original image to describe. Second, the dimension of Radon-pixel representation is much lower than that of the original image, which provides convenience for further feature extraction.

In summary, the Radon transform converts a pixel represented image into an equivalent, lower dimensional and more geometrically informative Radon-pixel image, which is a good basis of defining structural features for RR-IQA.

3.2. Fractal Analysis

Fractal analysis is introduced and developed by Mandelbrot [29] as a means for describing and analyzing the properties of objects with irregular and complex structure in nature. The characteristic property of fractals can be viewed as the objects with statistical self-similarity. The numerical quantification of self-similarity is obtained by the FD. The FD d is a measure of a given point set E in a certain measurement space $m(\cdot)$ by measuring its power law behavior with respect to the scale δ:

$$m_\delta(E) \propto \delta^{-d},$$

where $m_\delta(E)$ is some measurement of the given point set E at scale δ.

There are many techniques to estimate the FD of signal. One popular approach is the so-called *differential box counting* (DBC) method, which has the advantage of efficiency, accuracy and generality [30]. The DBC method considers a signal $f(x)$ of length M as a 2D point set $\{(x, y)|y = f(x)\}$, where x denotes the signal position and y denotes the magnitude level of the signal. Suppose the signal is scaled down to a size s, where s is an integer and $1 < s \leq M/2$. The x space is partitioned into grids of size s. A column of boxes of size $s \times s$ are placed on each grid respectively. Suppose the minimum magnitude value and the maximum magnitude value in the ith grid fall in the kth box and the lth

box respectively, we compute the contribution $n_r(i)$ in the ith grid as follows,

$$n_r(i) = l - k + 1. \tag{2}$$

Summing contributions from all grids, we have

$$N_r = \sum_i n_r(i), \tag{3}$$

where N_r is counted for different values of r. Then the DBC fractal dimension is defined as

$$d_{DBC} = \lim_{r \to 0} \frac{\log(N_r)}{-\log r}. \tag{4}$$

In practice, d_{DBC} can be estimated from the least squares linear fitting in the $log(N_r)$-$log(1/r)$ coordinates system.

In this work, fractal analysis is adopted to encode Radon-pixels image. The advantages of employed fractal analysis are as follows:

First, FD has a strong correlation with HVS [22]. Second, compared with statistical approaches, FD can encode spatial information in form of the geometrical distribution of the point sets[26].

As we known, intensity images of most natural surfaces are isotropic fractals [22]. To demonstrate that the image's projected response of Radon transform can also be characterized by the fractal model, we plot the behaviors of three projections by log-log fitting in Figure 2. It can be seen that the projections of the image do behave according to some power law.

4. Similarity of Radon-pixels Regularity(SRR)

Our proposed approach can be briefly described as the collection of FD defined on each projected component of Radon transform. The computation of the proposed approach is outlined in Algorithm 1. In the rest of this section, we will give detailed description of each step in the algorithm.

Algorithm 1 Description of SRRM

Input:

Reference image I_r;

Distorted image I_d;

Output:

$abs(SRR(I_r) - SRR(I_d)$;

1: Compute the Radon transform of I_r and I_d ,

$$\{R(I_r), R(I_d)\}; \tag{5}$$

2: Compute the FD for each projection in (5) using the box-counting method,

$$\begin{aligned} SRR(I_r) &= [FD_\theta(R(I_r))], \\ SRR(I_d) &= [FD_\theta(R(I_d))]; \end{aligned} \tag{6}$$

3: **return** $abs(SRR(I_r) - SRR(I_d)$.

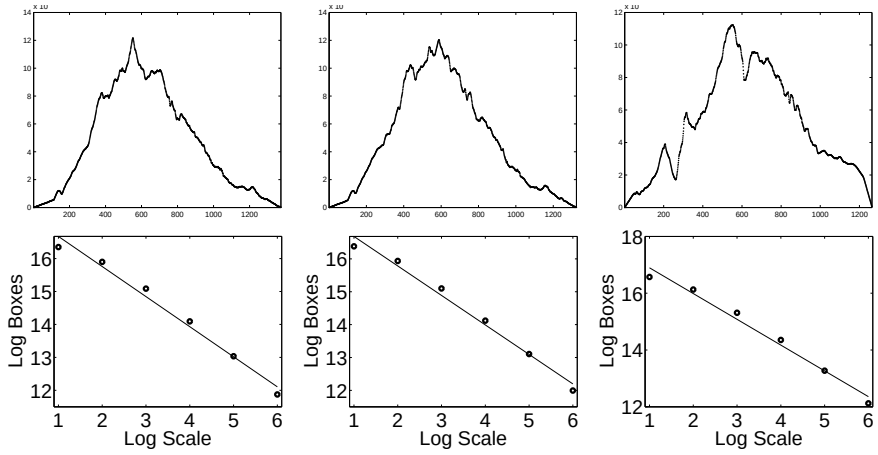

Figure 2: Log-log plot of box number versus box scale for three projections of the image by Radon transform. The upper row shows the projections of the reference image in Figure 1 in three different directions. The bottom row shows the corresponding log-log fittings.

4.1. Feature Based on Radon Transform

Given an image I, we propose the Radon transform for the image. Then we calculate the FD for each direction component of Radon transform using the box-counting method. Thus, we get the bag of FD for the given image I as follows,

$$SRR = \{FD_\theta(R_I)\}. \tag{7}$$

By using SRR, our approach can capture different structures of natural images from different directions. For each direction, we get one FD feature. Since 16 directions are used(It will be discussed in 5.2). Thus the image feature vector is 16 dimensional.

4.2. Similarity Index of SRR

Once the SRR features of the reference image I_r and the distorted image I_d have been obtained, we compute our SRR measure, denoted as *SRRM* by

calculating the ℓ_1 distance of the two feature vectors as follows,

$$SRRM(I_r, I_d) = ||SRR(I_r) - SRR(I_d)||_1. \tag{8}$$

Due to the nature of the SRR feature, the SRRM value measures the difference between the distorted image and the reference image in the meaning of the regularity of the projected spatial distribution of the image structures. It is noted that using ℓ_2 distance would yield similar performance in practice, and more complex distances [31–36], such as the Support Vector Machines(SVM) which has been widely demonstrated in different applications(e.g. image retrieval [36], image quality assessment [20]), might yield better performance. We employ ℓ_1 distance just for simplicity.

To verify the rationality of the SRRM method, we compute the SRRM of the reference image in Figure 1 with three most commonly encountered image distortions, which are blurring(with smoothing window), additive Gaussian noise(with zero-mean and the changing variance) and JPEG compression(with the changing compression rate). Meanwhile, we also compute the average SRRM values over all the reference images in the LIVEmd database with respect to above three types of distortion. The results are illustrated in Figure 3. In the light of the fact that an IQA metric can be viewed as a excellent metric as long as the metric monotonously changes with distortion increasing. What's more, from Figure 3, it can be seen that the prediction of the SRRM method trends to rise when the degree of the distortion increases. Thus, the SRRM values of our method are well consistent with the tendency of the decreasing image quality.

5. Experiment

We evaluated our SRR feature by applying it to the RR-IQA. In this section, we report on the performance evaluation and analyze the experimental results.

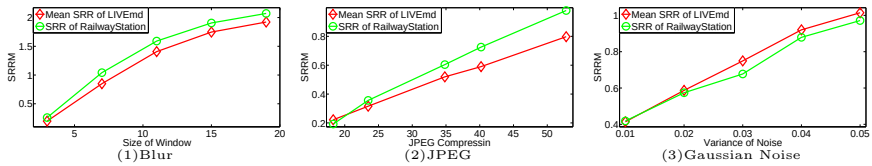

Figure 3: Rationality of SRRM. (1) SRRM of images with blured. (2) SRRM of images with JPEG compressed. (3) SRRM of images with Gaussian noise contaminated.

Table 1: Configurations of four benchmark datasets

Database	**Reference Img**	**Distorted Img**	**Distortion Types**
LIVEmd	15	450	2
LIVE	29	779	5
CSIQ	30	866	6
TID2013	25	3000	24

5.1. Benchmark Databases and Test Methodology

Subjective IQA studies are of fundamental importance for the development of IQA. Over the years, many researchers have contributed significant research in this area through the construction of various IQA databases. In this section, we introduce the composition of the four largest databases, i.e. the Tampere Image Database 2013(TID2013) [37], LIVE multiple distorted(LIVEmd) database [38], LIVE database [39], and Categorical Subjective Image Quality (CSIQ)database [40]. The important information of these four databases, in terms of the number of reference images, the number of distorted images, and the number of quality distortion types is summarized in Table 1.

The TID2013 is so far the largest image database for quality evaluation. It includes 3,000 images in sum. These images are generated by corrupting 25 original images with 24 types of distortion at 5 different levels. The

distortion types include: WN(#1), additive white Gaussian noise which is more intensive in color components than in the luminance component(#2), additive Gaussian spatially correlated noise(#3), masked noise(#4), high frequency noise(#5), impulse noise(#6), quantization noise(#7), Gblur(#8), image denoising (residual noise, #9), JPEG(#10), JP2k(#11), JPEG transmission errors(#12), JPEG2000 transmission errors(#13), non-eccentricity pattern noise(#14), local block-wise distortion of different intensity(#15), mean shift(#16), contrast change(#17), change of color saturation(#18), multiplicative Gaussian noise(#19), comfort noise(#20), lossy compression of noisy images(#21), image color quantization with dither(#22), chromatic aberrations(#23), and sparse sampling and reconstruction(#24). The MOS of each image is available and ranges from 0.2 to 7.3.

The LIVEmd database includes images distorted by multiple types of distortion. There are two subsets, one of which is associated with the images corrupted by Gblur followed by JPEG (GblurJPEG), and the other is associated with images corrupted by Gblur followed by WN (GblurWN). Each subset includes 225 images. The DMOS of each image is released and ranges from 0 to 85.

The CSIQ database consists of 866 images which are derived from 30 original images. Six types of distortion are considered in CSIQ, i.e. WN, JPEG, JP2k, additive Gaussian pink noise(PN), Gblur, and global contrast decrements(GCD). The DMOS of each image is available and ranges from 0 to 1.

The LIVE database includes 779 images, which are generated from 29 original images by corrupting them with five types of distortion, i.e. JPEG2000 compression(JP2k), JPEG compression(JPEG), WN, Gaussian blur(Gblur), and F-F. The DMOS and realigned DMOS of each image are available. Because the realigned DMOS is more precise than the DMOS, we adopt the realigned DMOS in our work. The realigned DMOS ranges from -3 to 112.

There are many criteria for quantifying the performance of the IQA ap-

proaches. Before computing these metrics, a regression analysis is required to provide a nonlinear mapping between the objective scores and the subjective mean opinion scores (MOS). In our setting, the following mapping function [41] is used,

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + exp(\beta_2(x - \beta_3))} \right) + \beta_4 x + \beta_5, \tag{9}$$

where $\beta_i, i = 1, 2, \ldots, 5$ are the parameters to be fitted by logistic regression. With the regression computed, five popular criteria are employed for evaluation in our experiment. These criteria are

- Pearson linear correlation coefficient (PLCC),
- Spearman rank-order correlation coefficient (SROCC),
- Kendall rank-order correlation coefficient (KROCC),
- Root mean square error (RMSE),
- Mean absolute error (MAE).

Roughly speaking, the PLCC, RMSE and MAE criteria are used to measure the prediction accuracy, while the SROCC and KROCC criteria are used to measure the monotonicity. All these evaluation criteria except KROCC are recommended by video quality experts group [42]. A desirable objective RR-IQA measure is expected to have high values for the SROCC, KROCC and PLCC criteria, and meanwhile have low values for the RMSE and MAE criteria.

5.2. Implementation of SRRM

We use the matlab function(radon(img, theta)) to implement the Radon transform. Thus, the only parameter in the proposed SRRM model is the projected number of the Radon transform, which determines the feature length of SRRM. We tuned the parameter based on TID2013, CSIQ and LIVE databases. Figure 4 plots the SROCC curves against numbers of projection by applying

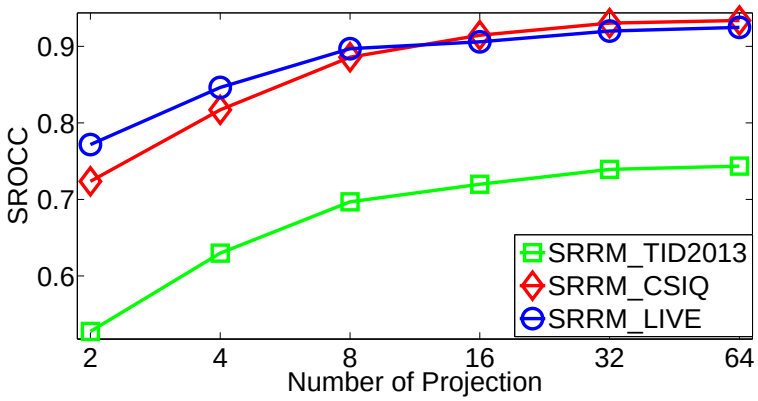

Figure 4: The performance of SRRM in terms of SROCC vs. the projected number of Radon transform on the TID2013, CSIQ and LIVE databases.

SRRM to three biggest databases. One can see that with the projected number increasing, the better results obtains. In our implementation, we set numbers of the projection 16 to get the balance between performance and data rate.

5.3. Performance Comparison

The experimental results of the proposed SRRM and the compared approaches on four benchmark databases are listed in Table 2 in terms of SROCC, PLCC, KROCC, RMSE and AME. In addition, in order to provide an overall indication of the comparative performance of the different schemes, Table 2 also gives the average PLCC, SROCC, and KROCC results over four databases, where the average values are computed in two cases [1]. In the first case, the correlation scores are directly averaged, whereas in the second case, different weights are assigned to the databases depending on the number of the distorted images in each database (refer to Table 1).

Generally speaking, from Table 2, we can see that with respect to the prediction performance, the proposed scheme performs consistently well across all the databases.

Specifically, compared with WNISM, the proposed SRRM outperforms it by a large margin no matter on individual database or on all four databases. One possible reason is that WNISM is a statistical metric which doesn't encode image's spatial information. Another possible reason is that WNISM performs well on the single distortion type but not very well on the whole database, as can be observed from the scatter plots with CSIQ database in Figure 5.

Compared with RR-SSIM, the superiority of SRRM is obvious on CSIQ database and they have similar performance on the LIVE database. The proposed scheme has a clear advantage in terms of data rate and time required (execution time will be compared in subsection 5.5).

Compared with RRED, SRRM has no advantage on performance. The reason is that RRED is local scheme metric while SRRM is global scheme metric. However, SRRM has a clear advantage in terms of data rate and time required (refer subsection 5.5).

Compared with SVD, one can see that the SRRM performs much better no matter which database is used and what the evaluation criterion is, although SVD uses much more information of the reference image. The reason lies that the singular values can not efficiently encode the structural information of images [20].

In order to further demonstrate the effectiveness of the proposed metric, SRRM is also compared with two classical FR-IQA metrics as follows:

- PSNR, which has a wide usage in the image processing literature. It also provides useful baseline comparisons.
- SSIM[18], which is state-of-the-art FR-IQA algorithm that has demonstrated competitive performance. It is also available online [43] that facilitates repeatable experimental verifications.

Compared with PSNR, one can see that the SRRM outperforms PSNR no matter which database is used and what the evaluation criterion is.

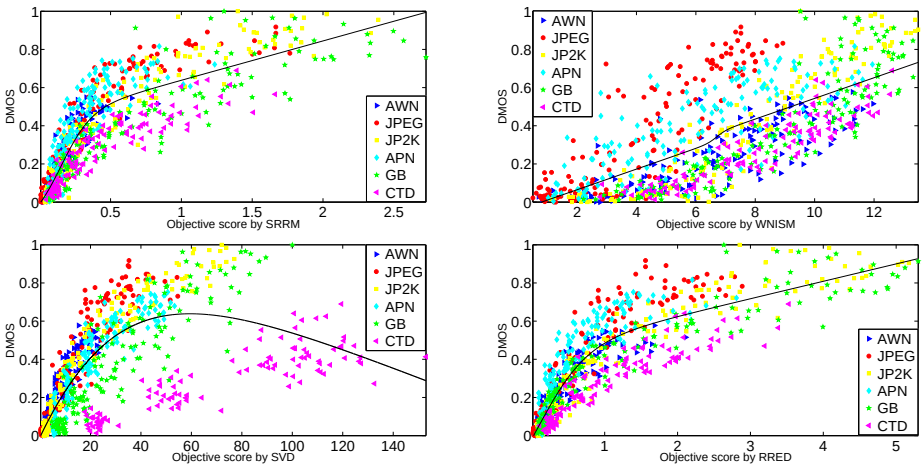

Figure 5: Scatter plots of predicted quality scores against the subjective quality scores (DMOS) by representative RR-IQA models on the CSIQ database. The six types of distortions are represented by different shaped colors.

Compared with SSIM, from the view of overall indication in Table 2, one can see that the SSIM has no apparent advantage than SRRM, and SRRM performs even better than SSIM in CSIQ database.

These further confirm that the proposed RR-IQA metric outperforms the state-of-the-art RR-IQA metrics.

In order to provide a visual illustration for the performance comparison among the compared RR-IQA models, scatter plots of subjective ratings versus objective scores on CSIQ database are shown in Figure 5, where each point represents one test image. The curves shown in Figure 5 are obtained by a nonlinear fitting according to (9). Compared with other scatter plots, points of SRRM and RRED are more close to each other, which means that SRRM and RRED correlate well with subjective ratings.

Table 2: Performance of the proposed SRRM and the other competing RR-IQA models in terms of SROCC, PLCC, KROCC, MAE, and RMSE on the LIVE, LIVEmd, CSIQ and TID2013 databases

DataBase	Criteria	SRRM	HWD2 [9]	WNISM [2]	RR-SSIM [1]	RRED [12]	SVD [19]	SSIM[18]	PSNR
Type		**RR**	**RR**	**RR**	**RR**	**RR**	**RR**	FR	FR
Length		**16**	**16**	**18**	**36**	**img.size/36**	**img.size/8**	img.size	img.size
TID2013	PLCC	0.7774	N/A	0.6247	N/A	0.8189	0.6715	0.7895	0.7018
	SROCC	0.7198	N/A	0.5202	N/A	0.7632	0.6534	0.7417	0.6394
	KROCC	0.5363	N/A	0.3647	N/A	0.5884	0.4855	0.5588	0.4696
	RMSE	0.7797	N/A	0.9680	N/A	0.7115	0.9186	0.7609	0.8832
	MAE	0.6059	N/A	0.7917	N/A	0.5407	0.7070	0.5927	0.6565
CSIQ	PLCC	0.9064	N/A	0.7124	0.8426	0.9121	0.8613	0.8613	0.8000
	SROCC	0.9146	N/A	0.7431	0.8527	0.9184	0.8756	0.8756	0.8058
	KROCC	0.7365	N/A	0.5457	0.6540	0.7429	0.6907	0.6907	0.6084
	RMSE	0.1109	N/A	0.1842	0.1413	0.1077	0.1334	0.1334	0.1575
	MAE	0.0866	N/A	0.1492	0.1092	0.0820	0.0991	0.0991	0.1195
LIVE	PLCC	0.9021	0.9624	0.7365	0.9194	0.9385	0.8266	0.9449	0.8723
	SROCC	0.9060	0.9418	0.7472	0.9129	0.9429	0.8251	0.9479	0.8756
	KROCC	0.7296	N/A	0.5577	0.7349	0.7888	0.6328	0.7963	0.6865
	RMSE	11.7887	6.3657	18.4814	11.3026	9.4317	15.3764	8.9455	13.3597
	MAE	9.0863	4.8445	14.6352	9.1889	7.2976	11.7104	6.9325	10.5093
LIVEmd	PLCC	0.8554	N/A	0.7347	N/A	0.9010	0.7388	0.8817	0.7315
	SROCC	0.8309	N/A	0.6217	N/A	0.8810	0.6223	0.8560	0.6709
	KROCC	0.6337	N/A	0.4412	N/A	0.6991	0.4504	0.6647	0.4945
	RMSE	6.4506	N/A	8.4484	N/A	5.4023	8.3916	5.8743	8.4914
	MAE	5.2191	N/A	6.6554	N/A	4.3018	6.6363	4.7662	6.8582
Weighted Average	PLCC	0.8253	N/A	0.6664	N/A	0.8603	0.7334	0.8336	0.7472
	SROCC	0.7912	N/A	0.6018	N/A	0.8275	0.7147	0.8061	0.7066
	KROCC	0.6085	N/A	0.4317	N/A	0.6551	0.5398	0.6269	0.5286
Direct Average	PLCC	0.8603	N/A	0.7021	N/A	0.8926	0.7745	0.8693	0.7764
	SROCC	0.8428	N/A	0.6581	N/A	0.8764	0.7441	0.8553	0.7479
	KROCC	0.6590	N/A	0.4773	N/A	0.7048	0.5648	0.6776	0.5647

5.4. Performance Comparison on Individual Distortion Types

To more comprehensively evaluate an IQA model's ability to predict image quality degradations caused by specific types of distortion, we compare the performance of competing methods on each type of distortion. The results are listed in Table 4, where only the SROCC scores are shown. There are a total of 37 groups of distorted images in the four databases.

We can directly view the performances of the compared metrics from Table 4. For the common distortion(see LIVE and CSIQ). HWD2 can accurately assess blocking artifact(JPEG) and is not fit for smoothing (GB). WNISM can accurately assess JP2K and is not fit for noise (AGN and PGN). RR-SSIM can accurately assess noise (AGN) and is not fit for smoothing (GB). SVD can accurately assess noise (AGN and PGN) and is not fit for smoothing (GB) and CTD. RRED and SRRM are mostly always better as a whole.

In order to demonstrate more clearly, in Table 3 we count the number of compared RR-IQA metrics on which the performances are better than SRRM according to SROCC values. One can see that SVD is counted 20 times in 37 groups. However, in above subsection, we have discussed that SRRM is much better than SVD no matter which database is used and what the evaluation criterion is. Generally speaking, performing well on specific types of distortions does not guarantee that an IQA model will perform well on the whole database with a broad spectrum of distortion types. A good IQA model should also predict the image quality consistently across different types of distortions. Referring to the scatter plots in Figure 5, it can be seen that the scatter plot of SRRM is more concentrated across different groups of distortion types. For example, its points corresponding to JPEG, PGN and CTD distortions are very close to each other. However, the points corresponding to JPEG, PGN and CTD for SVD are relatively far from each other. This explains why some RR-IQA models perform well for many individual types of distortions but they do not perform well on the entire database; that is, these IQA models behave

Table 3: The number of the performance, compared with RR-IQA models, is better than SRRM on each individual distortion type in terms of SROCC. Note, for HWD2 and RR-SSIM, the codes are not available, so we just compare with those we can get

Models	HWD2 [9]	WNISM [2]	RR-SSIM [1]	RRED [12]	SVD [19]
Length	**16**	**18**	**36**	**img.size/36**	**img.size/8**
SRRM	3/5	5/37	2/3	32/37	20/37

rather differently on different types of distortions, which can be attributed to the different ranges of quality scores for those distortion types.

5.5. Complexity

In applications such as real-time image quality monitoring and prediction, the complexity of implemented IQA models becomes crucial. We thus analyze the computational complexity of SRRM, and then compare the competing IQA models in terms of running time. Table 5 shows the running time of the seven IQA models on a pair of 1280×720 images which are taken from LIVEmd database. All algorithms were run on a Dell Inspiron INSP1440 notebook with Intel Core T6600 processor and 2GB RAM. The software platform used to run all algorithms was MATLAB R2012b. All the MATLAB source codes were obtained from the original authors.

Compared with FR-IQA metrics, SRRM takes more time than the PSNR and has the similar time with the SSIM.

Compared with RR-IQA metrics, clearly, SRRM is much faster than WNISM, RRED, SVD and RR-SSIM(according to [1], WNISM is about 2 times faster than RR-SSIM). The main reason is that WNISM, RRED, SVD and RR-SSIM take much time for the image decomposition while SRRM does not. Thus, the computational complexity of our approach is acceptable.

Table 4: Performance comparison of the IQA models on each individual distortion type in terms of SROCC

Database	Distortion	SRRM	HWD2 [9]	WNISM [2]	RR-SSIM [1]	RRED [12]	SVD [19]
Length		**16**	**16**	**18**	**36**	**img.size/36**	**img.size/8**
TID2013	#1	0.7644	N/A	0.6910	N/A	0.8519	0.9490
	#2	0.6980	N/A	0.6202	N/A	0.7893	0.8517
	#3	0.7875	N/A	0.6805	N/A	0.8517	0.9627
	#4	0.7736	N/A	0.6468	N/A	0.8042	0.7798
	#5	0.8521	N/A	0.7467	N/A	0.8961	0.9330
	#6	0.6688	N/A	0.6798	N/A	0.7931	0.7525
	#7	0.7488	N/A	0.6409	N/A	0.8282	0.8829
	#8	0.9263	N/A	0.9191	N/A	0.9666	0.8914
	#9	0.8911	N/A	0.8406	N/A	0.9233	0.9372
	#10	0.9055	N/A	0.8379	N/A	0.9274	0.9402
	#11	0.9408	N/A	0.9110	N/A	0.9539	0.9360
	#12	0.8151	N/A	0.8141	N/A	0.8468	0.7744
	#13	0.8241	N/A	0.7639	N/A	0.7897	0.8636
	#14	0.7126	N/A	0.4530	N/A	0.7809	0.8151
	#15	0.0260	N/A	0.2493	N/A	0.5491	0.5532
	#16	0.7581	N/A	0.4537	N/A	0.6143	0.6644
	#17	0.4015	N/A	0.5237	N/A	0.3970	0.3954
	#18	0.1814	N/A	0.1569	N/A	0.0455	0.0003
	#19	0.6923	N/A	0.6002	N/A	0.7950	0.8847
	#20	0.8476	N/A	0.7508	N/A	0.9039	0.9073
	#21	0.8489	N/A	0.7245	N/A	0.9203	0.9611
	#22	0.7813	N/A	0.4759	N/A	0.8646	0.8827
	#23	0.8781	N/A	0.8504	N/A	0.8915	0.8711
	#24	0.9180	N/A	0.9145	N/A	0.9536	0.9431
CSIQ	AGN	0.9282	N/A	0.8188	N/A	0.9353	0.9467
	JPEG	0.9614	N/A	0.8955	N/A	0.9521	0.9459
	JP2K	0.9596	N/A	0.9405	N/A	0.9628	0.9746
	APN	0.9166	N/A	0.8002	N/A	0.9362	0.9369
	GB	0.9467	N/A	0.9144	N/A	0.9634	0.9104
	CTD	0.9327	N/A	0.9122	N/A	0.9383	0.8651
LIVE	JP2K	0.9319	0.9362	0.9330	N/A	0.9580	0.9064
	JPEG	0.9524	0.9543	0.9204	N/A	0.9759	0.8821
	AGN	0.9578	0.9321	0.8701	0.9642	0.9780	0.9396
	GB	0.9523	0.8282	0.9145	0.8692	0.9675	0.6578
	FF	0.9108	0.9386	0.9227	0.9137	0.9427	0.8816
LIVEmd	GJPEG	0.8387	N/A	0.7085	N/A	0.8719	0.6951
	GWN	0.8168	N/A	0.5454	N/A	0.9005	0.5358

Table 5: Time cost of each metric

Model	SRRM	RRED [12]	RR-SSIM [1]	WNISM [2]	SVD [19]	SSIM[18]	PSNR
Type	**RR**	**RR**	**RR**	**RR**	**RR**	FR	FR
Time (seconds)	0.10863	3.853580	N/A	5.986130	2.425327	0.104198	0.016128

5.6. Discussions

The challenging task in the design of a RR-IQA metric is the selection of the smallest set of features that can support quality assessment effectively. These features should

- provide an efficient representation of the image;
- be sensitive to various image distortions;
- be relevant to the perceptual sensitivity of HVS.

Since the HVS is more sensitive to the structural information, in this paper, a RR-IQA is proposed using the regularity of Radon representation. Otherwise, different part of an image has the different contribution to humans perception, FD is used by SRRM to encode the spatial information of images. Moreover, from view of applications, SRRM is simplicity and low complexity, which make it more attractive and competitive to previous RR-IQA metrics.

6. Conclusion

In this paper, we propose an effective RR-IQA model, SRRM. In contrast to previous approaches, SRRM effectively extracts image structural information by Radon transform. Another feature of SRRM is that it effectively makes use of the spatial information by fractal analysis. Moreover, SRRM has only one parameter that does not need careful tuning. Finally, SRRM is computationally efficient. To demonstrate the power of the proposed approach, four largest

benchmark databases and five performance metrics are involved for evaluation. Our approach performs on a par with other state-of-the-art approaches. In the future, we will study the application of our approach to video quality assessment.

Actually, SRRM provides a general framework for RR-IQA. In this work, SRRM just uses the global features of a given image. In the future, we are seeking to further improve the current approach by using local image window for feature extraction.

[1] A. Rehman, Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation", IEEE Trans. Image Process. 21 (8) (2012) 3378–3389.

[2] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. H. Yang, A. C. Bovik, "Quality-aware images", IEEE Trans. Image Process. 15 (6) (2006) 1680–1689.

[3] H. R. Sheikh, A. C. Bovik, "Image information and visual quality", IEEE Trans. Image Process. 15 (2) (2006) 430–444.

[4] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, A. C. Bovik, "Image quality assessment based on a degradation model", IEEE Trans. Image Process. 9 (4) (2000) 636–650.

[5] L. Zhang, L. Zhang, X. Q. Mou, D. Zhang, "FSIM: a feature similarity index for image quality assessment", IEEE Trans. Image Process. 20 (8) (2011) 2378–2386.

[6] H. R. Sheikh, A. C. Bovik, L. K. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000", IEEE Trans. Image Process. 14 (11) (2005) 1918–1927.

[7] Z. Wang, A. C. Bovik, B. L. Evans, "Blind measurement of blocking arti-

facts in images", in: Proc. of IEEE Int. Conf. Image Process., Vol. 3, 2000, pp. 981–984.

[8] Z. Wang, H. R. Sheikh, A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images", in: Proc. of IEEE Int. Conf. Image Process., Vol. 1, 2002, pp. 477–480.

[9] X. Gao, W. Lu, D. Tao, X. Li, "Image quality assessment based on multiscale geometric analysis", IEEE Trans. Image Process. 18 (7) (2009) 1409–1423.

[10] U. Engelke, M. Kusuma, H. J. Zepernick, M. Caldera, "Reduced-reference metric design for objective perceptual quality assessment in wireless imaging", Signal Processing: Image Communication 24 (7) (2009) 525–547.

[11] Q. Li, Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation", IEEE Journal of Selected Topics in Signal Processing, Special issue on Visual Media Quality Assessment 3 (2) (2009) 202–211.

[12] R. Soundararajan, A. C. Bovik, "RRED indices: reduced reference entropic differencing for image quality assessment", IEEE Trans. Image Process. 21 (2) (2012) 517–526.

[13] Z. Wang, E. P. Simoncelli, "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model", in: Human Vision and Electronic Imaging X, Proc. SPIE, Vol. 5666, 2005, pp. 149–159, san Jose, CA.

[14] W. F. Xue, X. Q. Mou, "Reduced reference image quality assessment based on Weibull statistics", in: Proc. of the Int. Workshop on Quality of Multimedia Experience, 2010, pp. 1–6.

[15] L. Ma, S. N. Li, F. Zhang, K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation", IEEE Trans. on Multimedia 13 (4) (2011) 824–829.

[16] A. A. Abdelouahad, M. E. Hassouni, H. Cherifi, D. Aboutajdine, "Image quality assessment based on IMF coefficients modeling", in: Proc. of the Int. Conf. on Digital Information and Communication Technology and its Applications, 2011, pp. 131–145.

[17] I. P. Gunawanand, M. Ghanbari, "Reduced reference picture quality estimation by using local Harmonic amplitude information", in: Proc. of London Commun. Symp, 2003, pp. 137–140.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Trans. Image Process. 13 (4) (2004) 600–612.

[19] A. Shnayderman, A. Gusev, A. Eskiciogl, "An SVD-Based gray-scale image quality measure for local and global assessment", IEEE Trans. Image Process. 15 (2) (2006) 422–429.

[20] M. Narwaria, W. Lin, "SVD-Based quality metric for image and video using machine learning", IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics 42 (2) (2012) 347–364.

[21] S. R. Deans, "The Radon transform and some of its applications", Courier Dover Publications, 2007.

[22] A. P. Pentland, "Fractal-Based description of natural scenes", IEEE Trans. on Pattern Analysis and Machine Intelligence 6 (6) (1984) 661–674.

[23] Y. Quan, Y. Xu, Y. Sun, Y. Luo, "Lacunarity Analysis on Image Patterns for Texture Classification", in: CVPR2014., 2014, pp. 160–167.

[24] H. Ji, X. Yang, H. Ling, Y. Xu, "Wavelet domain multi-fractal analysis for static and dynamic texture classification", Transactions on Image Processing 22(1) (2013) 286–299.

[25] Y. Xu, X. Yang, H. Ling, H. Ji, "A New Texture Descriptor Using Multifractal Analysis in Multi-orientation Wavelet Pyramid", in: CVPR, 2010, pp. 161–168.

[26] Y. Xu, H. Ji, C. Fermüller, "Viewpoint invariant texture description using fractal analysis", International Journal of Computer Vision 83 (1) (2009) 85–100.

[27] Y. Xu, S. Huang, H. Ji, C. Fermuller, "Combining Powerful Local and Global Statistics for Texture Description", in: CVPR, 2009, pp. 573–580.

[28] T. M. Kusuma, H. J. Zepernick, "A reduced-reference perceptual quality metric for inservice image quality assessment", in: Proc. of Workshop on Mobile Future and Symposium on Trends in Communications, 2003, pp. 71–74.

[29] B. B. Mandelbrot, "The Fractal Geometry of Nature", San Francisco, CA: Freeman, 1982.

[30] N. Sarkar, B. B. Chaudhuri, "An efficient differential box-counting approach to compute fractal dimension of image", IEEE Trans. Systems, Man, Cybernet 24 (1) (1994) 115–120.

[31] J. Yu, M. Wang, D. Tao, "Semisupervised multiview distance metric learning for cartoon synthesis", IEEE Trans. Image Process. 21 (11) (2012) 4636–4648.

[32] J. Yu, D. Tao, J. Li, J. Cheng, "Semantic preserving distance metric learning and applications", Information Sciences, In Press.

[33] J. Yu, R. Hong, M. Wang, J. You, "Image clustering based on sparse patch alignment framework", Pattern Recognition, In Press.

[34] D. Tao, X. Li, X. Wu, S. Maybank, "Geometric mean for subspace selection", IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (2) (2009) 260–274.

[35] D. Tao, X. Li, X. Wu, S. Maybank, "General tensor discriminant analysis and gabor features for gait recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (10) (2007) 1700–1715.

[36] D. Tao, X. Tang, X. Li, X. Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (7) (2006) 1088–1099.

[37] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, et al., "Color image database TID2013: peculiarties and preliminary results ".

[38] D. Jayaraman, A. Mittal, A. K. Moorthy, A. C. Bovik, "Objective quality assessment of multiply distorted images", in: Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on, 2012, pp. 1693–1697.

[39] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z. Wang, A. C. Bovik, L. K. Cormack, "Image and video quality assessment research at LIVE", `http://live.ece.utexas.edu/research/quality`.

[40] E. C. Larson, D. M. Chandler, "Categorical Image Quality (CSIQ) Database", `http://vision.okstate.edu/csiq`.

[41] H. R. Sheikh, M. F. Sabir, A. C. Bovik, "A statistical evaluation of recent

full reference image quality assessment algorithms", IEEE Trans. Image Process. 15 (11) (2006) 3440–3451.

[42] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment", http://www.vqeg.org (2000).

[43] Z. Wang, "SSIM Index for Image Quality Assessment", https://ece.uwaterloo.ca/%7Ez70wang/research/ssim/.

Appendix

To make the paper more readable, a notation list is added as follows,

$\mathbf{R}_\theta(\rho)$ Radon transform, where θ is the angle and ρ is the signed distance to the origin of the coordinate system;

FD Fractal Dimension;

DBC Differential Box Counting;

$\mathbf{d_{DBC}}$ FD obtained by DBC;

SRR The bag of FD for the given image;

$\ell_\mathbf{1}$ L_1 norm;

SRRM $\ell_\mathbf{1}$ distance between the reference's SRR and the distorted's SRR.