Full-Reference Image Quality Assessment Approach Based on Image Separation

Bin Wang\textsuperscript{1,*}

\textsuperscript{1} Department of Mechanical and Electronic Engineering, Jiaxing University, Jiaxing, Zhejiang Province, China

\textsuperscript{a}iseewb@gmail.com

Keywords: Image Quality Assessment; Blind Source Separation.

Abstract. In this paper, a full-reference image quality assessment (IQA) approach based on blind source separation algorithm is present. After the reference image is divided into several patches, the scatter plot of these patches is contained within a simplicial cone. The mixing matrix is estimated by the parameter of the simplicial cone. The sources of the reference image and the distorted image are extracted by multiplying the image patches by the mixing matrix. Different sources of the reference image and the distorted image are used to calculate the structural similarity (SSIM) index values. SSIM values of different sources are weighted and summed to calculate the final objective IQA criterion. Experimental results show that the IQA approach outperforms the traditional single source IQA approach.

Introduction

Image quality assessment is an important and difficult problem in the area of image processing. It is widely used in video coding and computer vision. Over the past two decades, a lot of work has been done to develop the image quality assessment approaches. Those IQA approaches can be divided into two sorts, which are subjective IQA approach and objective IQA approach. Subjective IQA approach evaluates the image quality by human viewers, it is accurate and reliable. However, it is also time-consuming and cumbersome. Therefore, it needs to develop the algorithm which predicts the image quality automatically, that is the objective IQA approach. The objective IQA approach can be divided into three categories: full-reference IQA, reduced-reference IQA and no-reference IQA. Full-reference IQA evaluates the image quality by using the full information of the reference image. Reduced-reference IQA uses part of the information of the reference image. No-reference IQA do not use any information of the reference image at all. In this paper, we proposed a full-reference IQA approach by blind source separation algorithm.

Since the ultimate viewers of the image quality assessment are human beings, the most reliable IQA approach should be well correlated with the characteristics of human vision, that is the perceptual IQA approach. Many perceptual IQA approaches have been proposed to model the physiological properties of human vision system, among which the most well-known method is structural similarity (SSIM) index\cite{1}. SSIM extracts the structural information of the reference image and the distorted image and compares those informations to evaluate the image quality. The SSIM theory assumes that the main function of human vision system is to extract the structural information of the image and it achieves better performance than the traditional approaches such as MSE (Mean Square Error) and PSNR (Peak Signal Noise Ratio). Besides SSIM, other perceptual IQA approach derivated from SSIM are multiscale SSIM \cite{2}(MS-SSIM), information content-weighted SSIM \cite{3}(IW-SSIM) and complex wavelet SSIM\cite{4} (CW-SSIM). Video information fidelity \cite{5}(VIF) is another effective perceptual IQA approach. VIF is an IQA approach based on natural scene statistics, and it also has excellent correlation with the human perception of image quality. Video signal-to-noise ratio\cite{6} (VSNR) is a two-stage wavelet-based IQA approach for natural images, the contrast threshold are determined via a wavelet-based model in the first stage, and the distortions which are above the threshold are employed to compute the VIF value in the second stage.

This paper proposed a new image quality assessment approach based on blind source separation. The reference image and distorted image is divided into several patches, and patches of the reference
image or patches of the training image set are used to train the mixing matrix. After that, the patches of the reference image and the distorted image are separated into source patches by multiplying the mixing matrix, then the source patches of the reference image and the distorted image are compared by using the structural similarity index measurement. Finally, the SSIM of different source patches are pooled together to get the final IQA score.

The division of this paper is as follows: Section II describes the image blind source separation algorithm used in the paper. Section III presents the IQA approach by using the blind source separation. Section IV shows the experimental results of the IQA approach. Finally, conclusions are given in section V.

**Blind Source Separation Algorithm by Minimum Aperture Simplicial Cone**

The first step of our IQA approach is to separate patches of the reference image and the distorted image into source patches, it can be solved by the blind source separation process. The process of blind source separation of the image is divided into two steps: determining the mixing matrix and finding the source patches. At the beginning of the process of determining the mixing matrix, the reference image and the distorted image are separated into several image patches of sizes 8x8, each patch is transformed into column-wised vector of sizes 64x1. Those vectors are combined into sample matrix $X$. Given the observed sample matrix $X = \{x_1, x_2, \ldots, x_K\} = \{x_i | 1 \leq i \leq K\}$, the blind source separation (BSS) problem consists of estimating the source matrix $S = \{s_1, s_2, \ldots, s_K\} = \{s_i | 1 \leq i \leq K\}$ and the mixing matrix $A$. In which, $A \in R^{mn \times K}$, $x_i \in R^{64 \times 1}$, $s_i \in R^{64 \times 1}$ and $1 \leq i \leq K$. The blind source separation algorithm is given by:

$$X = AS$$

In which, $X \in R^{mn \times K}$ is the observation sample matrix, $A \in R^{mn \times K}$ is the mixing matrix, and $S \in R^{mn \times K}$ is the source matrix. Traditional BSS approach such as independent component analysis (ICA) assumes that the sources are independent of each other and it is not suitable for the dependent sources[11]. However, this is a much strict constraints for the BSS image sources. In fact, the independence between the sources is not necessary in the BSS problem of the image separation because the separate parts of the image may contain similar structure and contents. In this paper, we use the simplicial cone shrinking algorithm for unmixing non-negative sources [7] (SCSA-UNS) to separate the image patches. The mixing matrix and sources are estimated by finding the minimum aperture simplicial cone containing all of the mixed data points.

If the simplicial cone generated by the columns of mixing matrix $A$ is defined as[12]:

$$Span^+(A) = \{y | y = Ax, \text{ for } x \in R^n\}$$

Then the scatter plot of the mixed data $X$ is contained in the simplicial cone $Span^+(A)$. If the aperture of the simplicial cone is defined as:

$$\theta(Span^+(A)) = \frac{|\det(A)|}{\|A_1\| \times \|A_2\| \times \ldots \times \|A_n\|}$$

Then the minimum aperture simplicial cone is defined as the simplicial cone which contains the scatter plot of the mixed data and has the minimum aperture. If $Span^+(I_n)$ is the minimum aperture simplicial cone which contains the scatter plot of the sources, the mixing matrix $A$ can be estimated by seeking a minimum aperture simplicial cone which contains the scatter plot of the mixed data[7], in which, $I_n$ is the identity matrix, that is:

$$\hat{A} = \arg \min_{A \geq 0, A^+ \times x \geq 0} \theta(Span^+(A))$$

The mixing matrix can be estimated by the gradient ascent method. The flow of the algorithm of estimating the mixing matrix is described in paper[1].

525
Flow of Our Approach

The framework of our approach is shown in Fig. 1. The image set is divided into distorted images and original images. Since all image in the image set is color image, the color image is transformed into gray image at first, and then the distorted and original gray images are decomposed by blind source separation to n subimages $O_1, O_2, \ldots, O_n$ and $D_1, D_2, \ldots, D_n$ respectively. After that, the SSIM (structural similarity index) are calculated between $O_1$ and $D_1$, $O_2$ and $D_2$, ..., $O_n$ and $D_n$. The SSIM equation is used as follows:

$$C_i = \frac{2\mu(O_i)\mu(D_i) + C_1}{\mu(O_i)^2 + \mu(D_i)^2 + C_1} \cdot \frac{2\sigma(O_i, D_i) + C_2}{\sigma(O_i)^2 + \sigma(D_i)^2 + C_2}$$

(4)

In which, $\mu$ and $\sigma$ are mean value and standard deviation or covariance, $C_1$ and $C_2$ are two positive value avoided divisor to be zero. After $C_1, C_2, \ldots, C_n$ are calculated, the final metric is calculated as follows:

$$M = \sum_{i=1}^{n} \omega_i C_i$$

(5)

In which, $M$ is the objective IQA metric, and $\omega_i$ is the weighting parameter.

Implementation Result and Experiments

The LIVE image quality assessment database[6] is used in the experiments. The order of fractional Fourier transform is 0.25, 0.5, 0.75 and 1. The image is divided into 8x8 blocks. There are 29 reference images in the LIVE IQA database, and there are 779 distorted images. We divide 29 reference images into two parts: 20 images are used for training, and the remaining 9 images are used for testing.

The LIVE (Laboratory for Image and Video Engineering) image quality assessment database [3, 6, 7] is used to evaluate the performance of our approach. The LIVE IQA database has 29 reference image and 779 distorted image with five types, which are JPEG2000 Compression (JPEG2K), JPEG Compression (JPEG), White Noise (WN), Gaussian Blur (GB) and Fast Fading (FF). The LIVE IQA database is divided into two parts: training image and testing image. The degraded version of 24 reference images are used for training, and the degraded version of remaining 5 reference images are used for testing.

The Spearman rank order correlation coefficient (SRCC) and Pearson linear correlation coefficient (PLCC) are used to evaluate the performance of different algorithms. Two other approaches which are PSNR and MS-SSIM are used to compare with our approach. The detailed result is shown in Table I. According to table I, the proposed approach is superior than the PSNR approach by evaluating the SRCC and PLCC coefficients between the distorted image and the reference image in testing image database.
Table I. Spearman Rank Order Correlation Coefficient of Different Algorithm (aligned DMOS)

<table>
<thead>
<tr>
<th></th>
<th>JP2K</th>
<th>JPEG</th>
<th>WN</th>
<th>GB</th>
<th>FF</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>0.8954</td>
<td>0.8809</td>
<td>0.9854</td>
<td>0.7823</td>
<td>0.8907</td>
<td>0.8755</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.9613</td>
<td>0.9754</td>
<td>0.9748</td>
<td>0.9474</td>
<td>0.9537</td>
<td>0.9478</td>
</tr>
<tr>
<td>VIF</td>
<td>0.9322</td>
<td>0.9533</td>
<td>0.9531</td>
<td>0.9433</td>
<td>0.9532</td>
<td>0.9472</td>
</tr>
<tr>
<td>Proposed</td>
<td>0.9157</td>
<td>0.9259</td>
<td>0.9632</td>
<td>0.8528</td>
<td>0.9173</td>
<td>0.9121</td>
</tr>
</tbody>
</table>

Table II. Pearson Correlation Coefficient of Different Algorithm (aligned DMOS)

<table>
<thead>
<tr>
<th></th>
<th>JP2K</th>
<th>JPEG</th>
<th>WN</th>
<th>GB</th>
<th>FF</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>0.8769</td>
<td>0.8674</td>
<td>0.9294</td>
<td>0.7746</td>
<td>0.8760</td>
<td>0.8591</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.8845</td>
<td>0.9268</td>
<td>0.9560</td>
<td>0.8820</td>
<td>0.9349</td>
<td>0.8959</td>
</tr>
<tr>
<td>VIF</td>
<td>0.8833</td>
<td>0.9425</td>
<td>0.9326</td>
<td>0.9325</td>
<td>0.8532</td>
<td>0.8942</td>
</tr>
<tr>
<td>Proposed</td>
<td>0.8942</td>
<td>0.9031</td>
<td>0.9432</td>
<td>0.8231</td>
<td>0.8873</td>
<td>0.8783</td>
</tr>
</tbody>
</table>

Conclusions

In this paper, the blind source separation of images are used to predict the quality score of the distorted image. The distorted and original images are processed by blind source separation at first, then the image quality of decomposed component of the distorted image are evaluated by using SSIM metric. Finally, SSIM of every components are pooled together to get the objective image quality score. The experiment result of LIVE IQA database shows that our approach can compete with the traditional algorithm in evaluating the quality score of the image.

References