Computers and Electrical Engineering 000 (2016) 1–17

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Image quality assessment via spatial structural analysis

Xichen Yang*, Quansen Sun, Tianshu Wang

School of Computer science and engineering, Nanjing University of Science and Technology, Xiaolingwei 200, Nanjing 210094, China

A R T I C L E I N F O

Article history:
Received 9 March 2016
Revised 18 August 2016
Accepted 18 August 2016
Available online xxx

Keywords:
Full-reference
Image quality assessment
Distortion
Spatial structural distribution
Structural similarity

A B S T R A C T

The human visual system is sensitive to image structural information. Modeling of image structural similarity has been regarded as suitable for achieving perceptual quality predictions. However, most structural similarity-based image quality assessment (IQA) methods focus on spatial contrast without fully considering the spatial structural distribution. Hence, we propose an IQA method that considers both spatial contrast and structural distributions. First, the image gray-scale fluctuation map (GFM) is calculated. Second, the spatial structural information variation matrices (SSVMs) between the GFMs of distorted and pristine images are obtained. Finally, the quality prediction model is trained using support vector regression (SVR). The experimental results show that the proposed method can accurately predict human perceptual image quality. Experiments on the LIVE2 database show that the Spearman rank-order correlation coefficient (SROCC) and linear correlation coefficient (LCC) values exceed 0.85, while the scale or distortion type of the training set changes, which indicates stability.

1. Introduction

We are currently living in a multimedia era, wherein a great variety of new handheld devices and smart phones has emerged. Consequently, perceptual optimizations of multimedia services have gained an important position. Since images are one of the most important carriers of visual information interaction [1], image quality assessment (IQA) plays an increasingly important role in multimedia service performance comparisons [2]. IQA methods can be divided into two categories: subjective assessment by human eyes and objective assessment by automatic quality assessment algorithms . Since the human eye is the receiving terminal of visual information, subjective assessment is the ultimate criterion of image visual quality. However, subjective assessment is time-consuming and expensive, and the assessment results are vulnerable to certain factors, e.g., the educational background and working environment of the reviewers. Hence, objective assessment is more suitable for practical applications. Hence, finding an objective IQA that can automatically predict image quality in a manner that agrees with human vision is a crucial step for any image processing system.

According to the degree of information needed for IQA, objective IQA can be divided into three categories: (1) Full-reference (FR) IQA [1,3–10], (2) Reduced-reference (RR) IQA [11,12], and (3) No-reference (NR) IQA [13,14]. The most direct and accurate way to measure service quality is to compare the distortion degree of the images provided from different service systems, which contain the same scene contents and the same corresponding pristine images. Since multimedia service quality performance comparisons are mostly taken by service providers or professional testers, image data of certain requirements are available. Hence, FR IQA is most suitable for multimedia service quality performance comparisons.

* Corresponding author.
E-mail address: yxcwudi@gmail.com (X. Yang).

http://dx.doi.org/10.1016/j.compeleceng.2016.08.014

In general, the idea of FR IQA is to predict image quality by measuring the similarity or difference between the distorted image and its corresponding pristine image. The peak signal-to-noise ratio (PSNR) metric is the most popular FR IQA [3], which directly measures the differences between the distorted image and the reference image by calculating pixel-by-pixel information. Though the PSNR is well-defined with a clear physical meaning, it does not consistently coincide with human visual perception. Therefore, a considerable number of efforts have been made to develop IQA methods that can mimic the subjective assessment procedure of the human visual system (HVS). Among all of the HVS-oriented IQA methods, the structural similarity-based quality metric (SSIM) [4] is the most accepted method. Based on the assumption that the HVS is highly adapted to extract structural information for image visual content understanding, the SSIM metric predicts image quality by measuring structural degradation. In addition to SSIM, there are many other HVS-oriented IQA methods, such as the multiscale SSIM (MSSIM) [5], which extends the SSIM into multiscale space. In [6], the feature similarity index (FSIM) is proposed by using phase congruency and the image gradient magnitude to characterize local structure degradation. In [7], gradients are used to capture image structural information. In addition to these structure similarity-based methods, many other interesting works have also been proposed for IQA such as the noise quality measure (NQM) [8] and information fidelity criterion (IFC) [9]; IFC was further expanded to the visual information fidelity (VIF) in [1]. These HVS-oriented IQA methods promote our understanding of perceptual quality assessment.

The HVS is sensitive to image structural information degradation. The most straightforward understanding of image structure is given by the relationships among pixels in the image area [15], which include both spatial contrast and the spatial structural distribution. Existing structural similarity-based IQA metrics [4,5,7] mainly consider the influence of distortion for spatial contrast; changes in the spatial structural distribution is not fully considered. Hence, we take the spatial structural distribution into consideration. How to describe the image spatial structure is still an open problem. An image can be regarded as a three-dimensional space, and the surface structure in different regions can be expressed by gray-scale values. By analyzing gray-scale fluctuations in different regions of the image, we can learn more about the spatial structural distributions of the joint pixels. In this paper, we use image gray-scale fluctuations to describe spatial structural information. Moreover, image distortions may cause changes in the structural information in the same regions between the reference image and the distorted image, e.g., blur distortions may change a region that has a relatively large gray-scale fluctuation into a flat region. Hence, image gray-scale fluctuation is used to measure the degradation on the spatial structural distribution.

In this paper, we propose an FR IQA method to measure image structure information changes caused by distortion in both structural intensity and structural distributions. First, a specific image primitive is introduced to analyze image gray-scale fluctuations [16], and the gray-scale fluctuation primitive map (GFM) of each image is obtained. The GFM represents the structural distribution of the corresponding image, and the gray-scale fluctuation shift between the GFMs of the distorted and reference images are defined to express structural distribution changes. Second, the structural intensity degradation is calculated as the contrast change between the GFMs of the distorted and reference images. Then, for each pair of gray-scale fluctuation shifts, the corresponding structural intensity degradation is cumulated and defined as the spatial structural information variation matrix (SSVM). Finally, SVR is employed to train the image quality prediction model [17], where the two inputs are SSVMs of the training set and human opinion scores. In this paper, we conduct comparison experiments on four open databases: LIVE2 [18], CSIQ [19], TID2008 [20], and LIVE Multiply Distorted [21]. The experimental results demonstrate that the new IQA method proposed in this paper has a high correlation with human subjective judgments of diversely distorted images.

The remainder of this paper is organized as follows. Section 2 presents the process of image structural information extraction. Section 3 presents the detailed procedure of the new IQA method. Section 4 discusses the results of our comparison experiments. Finally, in Section 5, we conclude our paper.

2. Image structural feature

Image distortion can immediately lead to changes in image spatial structural information. Gray-scale fluctuations of joint pixels in specific regions represent the shape characteristics of the image structure; for example, in edge regions, gray-scale fluctuations may be relatively large. On the contrary, smooth regions mostly correspond to small gray-scale fluctuations. Thus, in this paper, the image gray-scale fluctuation is used as the basis for analyzing image spatial structural information.

2.1. Image gray-scale fluctuation primitive

Image gray-scale fluctuations can be analyzed using a specific primitive that was introduced in [16]. The gray-scale relationships between a certain pixel and its neighbors in different detection directions are used to represent the gray-scale fluctuation of a pixel. Thus, the size of the primitive is set as 3×3. As shown in Fig. 1, the primitive is a 3×3 square, and pixels in the square are denoted by V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, and V_9. The detection directions are set to 0°, 45°, 90°, and 135°. In this way, the primitive can be used to obtain the gray-scale relationships between the center pixel and all of its neighbors.

To measure image gray-scale fluctuations, we define two variables: the neighbor pixel gray-scale vector angle and neighbor gray-scale mutually exclusive value. The two variables are denoted by Ga_x and Dt_x $\{x = 1, 2, 3, 4\}$, respectively, where x represents the detection direction (0°, 45°, 90°, and 135°). The calculation method for Ga_x is shown in Fig. 2. The central pixel of the primitive is defined as the origin of the coordinate axes (o). The neighbors of the central pixel in the current