Signal Processing: Image Communication 45 (2016) 1–9

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Multiscale contrast similarity deviation: An effective and efficient index for perceptual image quality assessment ☆

Tonghan Wang [a,*], Lu Zhang [b], Huizhen Jia [c], Baosheng Li [a,d], Huazhong Shu [a]

[a] Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu, P.R. China
[b] IETR lab (UMR CNRS 6164), INSA de Rennes, 20 Avenue des Buttes de Coesmes, CS 70839F, 35708 Rennes Cedex 7, France
[c] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu, P.R. China
[d] Shandong Cancer Hospital, Ji-Nan, Shandong, P.R. China

ARTICLE INFO

Article history:
Received 24 November 2015
Received in revised form
20 April 2016
Accepted 20 April 2016
Available online 21 April 2016

Keywords:
Contrast similarity
Image quality assessment
Multiscale
Standard deviation pooling
Full reference

ABSTRACT

Perceptual image quality assessment (IQA) uses a computational model to assess the image quality in a fashion consistent with human opinions. A good IQA model should consider both the effectiveness and efficiency. To meet this need, a new model called multiscale contrast similarity deviation (MCSD) is developed in this paper. Contrast is a distinctive visual attribute closely related to the quality of an image. To further explore the contrast features, we resort to the multiscale representation. Although the contrast and the multiscale representation have already been used by other IQA indices, few have reached the goals of effectiveness and efficiency simultaneously. We compared our method with other state-of-the-art methods using six well-known databases. The experimental results showed that the proposed method yielded the best performance in terms of correlation with human judgments. Furthermore, it is also efficient when compared with other competing IQA models.

1. Introduction

Image quality assessment occupies a very important position in numerous fields and applications, such as image acquisition, compression, transmission and restoration. Since human beings are the ultimate receivers of the visual stimulus, it is essential to develop a perceptual model to closely correlate with the human visual system (HVS).

Objective quality assessment methods can be classified into three types [1]: (1) full-reference(FR), where an ideal "reference" image is available for comparison; (2) reduced-reference (RR), where partial information about the reference image is available; and (3) no-reference (NR), where the reference image is not accessible. This paper focuses on the FR methods. In the past decades, great efforts and huge advances have been made in FR methods. Here we briefly review some representative ones. More comprehensive surveys on FR-IQA metrics can be found in [16,17] and [18]. The traditional metrics such as the peak signal-to-noise ratio (PSNR) and the mean squared error (MSE) did not correlate well with human opinions [2]. The later developed FR methods could be generally classified into three types of approaches: the HVS model based ones, the information theoretic ones and the structural ones.

The noise quality measure index (NQM) [7], the visual signal-to-noise ratio index (VSNR) [8] and most apparent distortion (MAD) [12] are the three representatives HVS based FR methods. The NQM and the VSNR quantified the effects of different visual signals (e.g. the luminance, the contrast, the frequency content, the interaction between them) on the HVS. The MAD was proposed by Larson and Chandler based on the hypothesis that the HVS used different strategies for high quality and low quality images. However this kind of FR methods is usually not computationally efficient.

The information theoretic approaches include the visual information fidelity (VIF) [9] and the information fidelity criteria (IFC) [10]. The VIF took the FR IQA problem as an information fidelity problem and chose the amount of information shared by the reference image and the distorted one as the similarity, which was an extended version of the IFC.

The structural approaches are based on the assumption that the HVS is highly adapted for extracting structural information from the visual scene. As a milestone in the development of IQA models, the structural similarity (SSIM) [3] surpassed the previous ones since it showed a better correlation with the human perception. Then a lot of SSIM-based metrics have been proposed in the literature [4–6]. In [11], Gao et al. presented a content-based metric

☆The MATLAB source code of MCSD is public available online at http://www.mathworks.com/matlabcentral/fileexchange/56604-multiscale-contrast-similarity-deviation/content/mcsd.m

* Corresponding author.
E-mail address: thwang_seu@163.com (T. Wang).

http://dx.doi.org/10.1016/j.image.2016.04.005
0923-5965/

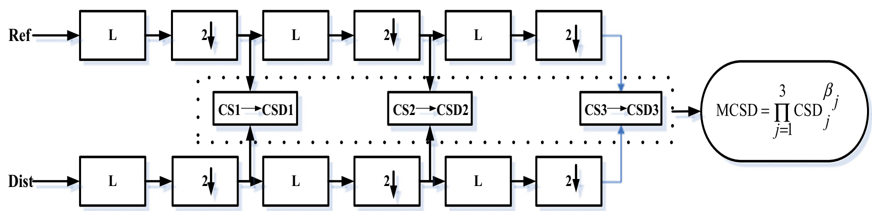

Fig. 1. Structure diagram of the MCSD. Ref indicates a reference image and Dist indicates the corresponding distorted one, L: low-pass filtering; 2↓: down sampling by 2; CS1, CS2 and CS3 mean contrast similarity map1, contrast similarity map2 and contrast similarity map3, respectively; CSD1, CSD2 and CSD3 mean the standard deviation of CS1, CS2 and CS3, respectively. The general idea of the MCSD is to decompose the reference and the distorted images into three scales, then calculate their contrast similarity deviation for each scale, and get the final note via a pooling of the CSD values on the three scales.

(CBM), which divided the structural information into edges, textures and flat regions in accordance with the content and then pooled the three parts with different weights to obtain the final image quality. Based on the observation that the visual information in an image is often redundant and the HVS understands an image mainly based on its low-level features, Zhang et al. proposed the feature-similarity (FSIM) index [14] which employed two features (the phase congruency and the gradient magnitude) to compute the local similarity map. Considering the gradients' sensitivity to structure and contrast changes, Liu et al. proposed GSM [13] based on a gradient similarity. Unlike the SSIM's average pooling, the CBM, the FSIM and the GSM adopted a weighting strategy for the pooling. Note that the weighting pooling may gain more IQA accuracy than those with average pooling to some degree, but it may increase the computational complexity. In addition, this pooling could make the predicted quality scores nonlinear to human opinions [15]. Based on these considerations, Xue et al. [15] proposed a gradient based model, based on the observation that the image gradient can effectively capture image local structures to which the HVS is highly sensitive. The gradient magnitude similarity deviation (GMSD) index [15] firstly the generated image gradient magnitude maps of the reference image and the distorted one, and then computed the similarity map of them, finally took the standard deviation of similarity map as the overall image quality score.

Generally, the effectiveness, namely, high correlation with the human subjective score, is the prerequisite of a good IQA model. The efficiency (low computation cost), however, is the second most important criterion of a good IQA model. Thus the effectiveness and efficiency are two ultimate goals for the design of IQA models, but unfortunately it is hard to reach these two goals simultaneously. Among the above-mentioned methods, the GMSD [15] had a big success in terms of the two goals, but its performances were a little bit low for certain distortion types (such as the contrast distortion). In this paper, we attempt to make another effort to fill this need and to overcome the problem of the GMSD by proposing an effective and efficient FR-IQA model called multiscale contrast similarity deviation (MCSD). The MCSD is also a structural approach.

The rest of this paper is organized as follows. In Section 2, we present the proposed model in detail. Section 3 shows and discusses the results. We conclude this work in Section 4.

2. Multiscale contrast similarity deviation (MCSD)

Contrast is a good attribute for characterizing the quality of an image [19]. Proper contrast change may even improve the perceptual quality of images. In fact, we can define “high quality” as appropriate contrast and little distortion. The contrast has been widely used in the area of image enhancement [20–25]. Contrast has previously been utilized in SSIM [3], where it was used as one of the three features – luminance, contrast and structure. Here we use the contrast feature alone to design our IQA model. Furthermore, the contrast is sensitive to the spatiotemporal frequency and viewing distance [26,27], which are related to the multi-scale representation to some extent. In fact, the multi-scale method is a convenient way to incorporate image details at different resolutions. Perceptibility of image details depends on viewing distance and sampling density of an image. Furthermore, a natural image might have objects and structures that are relevant at different scales, but the human eye is readily able to identify and process the information presented by it [28]. Thus, processing an image at various scales adds flexibility to the processing technique, and image scales play a very important role in IQA [4,29]. In [4], five scales, namely, the original image scale plus the other four reduced resolution (each down sampled by a factor of 2), were utilized to design the multi-scale SSIM. Very recently, by using proper image scales, the authors of [29] designed a totally training free NR-IQA. For these reasons, we combined the contrast with the multi-scale representation to design our model.

2.1. Structure diagram of the MCSD

The structure diagram of the MCSD is shown in Fig. 1. The design choices (all the parameters' settings) for MCSD will be introduced in Section 3.2.

The MCSD explores the contrast features by resorting to the multi-scale representation. The reason is that multiscale method incorporates image details at different resolutions, and contrast is relevant to the viewing distance. The contrast combined with multiscale representation were widely used in the literature of image enhancement [32–34,38,39] and image fusion [35–37], where the multiscale representation was implemented by pyramid decomposition [34,36], wavelet decompositions [32,34,36], scale-space approach [33] and homomorphic transform [39]. In view of efficiency, the multiscale representation in the proposed model, however, was operated by iteratively applying a low-pass filter and downsampling the filtered image by a factor of 2, as did in MS-SSIM [4].

Considering the computational cost, we do not use the original image scale but three reduced resolutions (each down sampled by a factor of 2). For the reference image and distorted image at each scale, we calculate their contrast similarity deviation (CSD), sort of contrast similarity map between them (the derivation of the CSD will be detailed in the next subsection). Then the CSDs at three scales are pooled to have the final score MCSD using the following equation:

$$MCSD = \mathop{\Pi}_{j=1}^{\text{nscales}} CSD_j^{\beta_j} \tag{1}$$

where Π means product, n scales is the total scale we use, j is the jth scale, $\beta_j$is the corresponding weight of the jth scale and $\sum_{j=1}^{\text{nscales}} \beta_j = 1$.