Neurocomputing 000 (2018) 1–11

Contents lists available at ScienceDirect

Neurocomputing

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

Referenceless quality metric of multiply-distorted images based on structural degradation

Tao Dai[a], Ke Gu[b], Li Niu[c], Yong-bing Zhang[a], Weizhi Lu[a,*], Shu-Tao Xia[a]

[a] Graduate School at Shenzhen, Tsinghua University, Guangdong, Shenzhen 518055, China
[b] Faculty of Information Technology,Beijing University of Technology, Beijing 100124, China
[c] Electric and Computer Engineering department, Rice University, Houston, TX 77005, USA

ARTICLE INFO

Article history:
Received 17 June 2017
Revised 27 December 2017
Accepted 11 February 2018
Available online xxx

Communicated by X. Gao

Keywords:
Image quality assessment
No-reference
Local binary pattern
Multiple distortions
Structural degradation

ABSTRACT

Multiply-distorted images, that is, distorted by different types of distortions simultaneously, are so common in real applications. This kind of images contain multiple overlaying stages (e.g., acquisition, compression and transmission stage). Each stage will introduce a certain type of distortion, for example, sensor noise in acquisition stage and compression artifacts in compression stage. However, most current blind/no-reference image quality assessment (NR-IQA) methods are specifically designed for singly-distorted images, thus resulting in their deficiency in handling multiply-distorted images. Motivated by the hypothesis that human visual system (HVS) is adapted to the structural information in images, we attempt to assess multiply-distorted images based on structural degradation. To this end, we use both first- and high-order image structures to design a novel referenceless quality metric for multiply-distorted images. Specifically, we leverage the quality-aware features extracted from both the gradient-magnitude map and contrast-normalized map, and further improve the performance by making use of redundancy of features with random subspace method. Experimental results on popular multiply-distorted image databases verify the outstanding performance of the proposed method.

1. Introduction

Recent years have witnessed a fast development of imaging devices, such as digital cameras, and smart phones. The fast development of imaging devices has put forward the requirement for high quality of experience of end-users [1]. In practice, multiply-distorted images, namely images with multiple distortion types, are common in real applications. The different types of distortions often arise from overlaying stages, such as signal acquisition, compression and transmission stages. The appearance of multiple distortions makes the problem of blind/no-reference image quality assessment (NR-IQA) even more complex, since the modeling of the interaction between the single distortions and their interactive effects on the overall visual quality remains still challenging [2]. Besides, most existing NR-IQA methods are designed for singly-distorted images and struggle to achieve satisfying results on multiply-distorted images [3–5]. Thus, it is still demanding to develop new NR-IQA methods for multiply-distorted images.

Many early works attempt to assess the perceived quality of distorted images by considering structural information of an image. It has been widely acknowledged that structural information, like the image gradient, plays a key role in evaluating perceptual quality [6–8]. To date, there have existed many successful IQA methods based on structural degradation [4,9–11]. Taking the successful structural similarity index (SSIM) and its variants [9,10] for example, these structural-based methods achieve remarkable performance by exploring the structural information. The performance advantage comes from the fact that human visual system (HVS) is highly adapted to extract the structural information from a scene. In general, it would have two main advantages based on structural degradation to evaluate image quality: 1) it is an effective way to measure the degree of distortion based on structural degradation; 2) structural degradation contributes to distinguishing different types of image distortions. In other words, structural degradation intuitively reflects the overall quality of the perceived image.

Inspired by the above observations, some recent NR-IQA methods [4,11] have also been developed for multiply-distorted images based on the potential structural degradation, given that the image structure conveys essential visual information from a scene. In these methods, local binary pattern (LBP) operators are applied to extract structural features from different domains, such as the gradient domain [4] and the perceptual opponent-color domain [11]. Unfortunately, these methods do not fully exploit the image structural information. More precisely, most existing

* Corresponding author.
E-mail addresses: dait14@mails.tsinghua.edu.cn (T. Dai), guke.doctor@gmail.com (K. Gu), wzlusd@sz.tsinghua.edu.cn (W. Lu), xiast@sz.tsinghua.edu.cn (S.-T. Xia).

https://doi.org/10.1016/j.neucom.2018.02.050

methods only utilize first-order image structures (e.g., gradient information) without taking account of high-order image structures (e.g., texture information). In practice, however, high-order image structures also play a vital role in evaluating the visual quality, due to the fact that HVS has separate mechanisms to process the first- and high-order image structures [12].

To make use of various types of image structures better, we develop a novel blind/NR-IQA metric of multiply-distorted images based on structural degradation, which considers both the first- and high-order image structures. In fact, the first-order image structures characterize the main structures of images (e.g., edges), while the high-order image structures represent fine-grained details of images (e.g., textures) [12]. Such two structural features are sensitive to different types of distortions, which thus renders them effective to reflect the perceptual quality of images. In order to utilize such image structures better, we attempt to extract a number of quality-aware features from the gradient magnitude map and contrast-normalized map, which represent the first- and high-order structural patterns of distorted images, respectively. Specially, the rotation invariant uniform LBP operator is adopted to extract the first-order structural patterns, while a generalized center-symmetric LBP (GCS-LBP) operator is employed to extract high-order structural patterns. For clarity, the main contributions of the work are summarized as below:

- Based on structural degradation, we utilize both the first- and high-order image structures to characterize the image structural features.
- Technically, when extracting the first-order structural information, we employ four directional high-pass filters (5×5) encompassing the diagonal directions to incorporate neighboring information. When extracting the high-order information, we propose a generalized center-symmetric LBP to resist the visual artifacts introduced by the contrast modifications.
- In order to avoid overfitting, we apply the random subspace method [13] in the feature space. Experimental results confirm that our proposed method acquires superior performance, by comparison with prevailing full reference (FR) and no reference IQA methods.

The rest of this paper is constructed as follows. In Section 2, we briefly review some related work on objective IQA metrics. In Section 3, we introduce the proposed IQA metric in detail, and analyze the effects of feature set. Then, experimental results are provided in Section 4. Finally, we draw the conclusions in Section 5.

2. Related work

Objective IQA algorithms aim to evaluate image visual quality automatically according to computational models, which have been successfully and widely applied in the field of image processing, such as image enhancement [14,15], image denoising [16–18], and image classification [19,20]. According to the amount of available information from the pristine image (reference image), existing objective visual quality metrics can be roughly divided into full-reference (FR), reduced-reference (RR) and no-reference (NR) IQA metrics. Most early IQA methods belong to the class of FR-IQA methods. Traditional FR-IQA metrics like peak signal-to-noise ratio (PSNR) and mean square error (MSE) are the most commonly-used IQA metric in different applications for their simplicity and efficiency. Nevertheless, it is shown that both PSNR and MSE can not coincide with human subjective ratings very well [1,21]. To overcome this drawback, more advanced FR-IQA methods that highly agree with human perceptual quality have been proposed, such as SSIM [22], OSS-SSIM [10], VIF [23], MAD [24], ADM [25], FSIM [6], GMS [7], and GMSD [8]. RR-IQA algorithms [26,27] tend to compare the partial information between the distorted image and the corresponding reference image. However, in most practical applications (e.g., denoising, enhancement), the reference image is generally unavailable, thus constraining the usage of FR- and RR-IQA methods. Hence, it is of high demand to design NR-IQA methods in practice, without access to reference image.

Most of early NR-IQA metrics have been specifically designed to reflect the visual quality of images in particular applications, like Gaussian blur (GB) [28,29], white noise (WN) [30], JPEG compression [29,31,32], and contrast distortion [33]. Thus, these methods can be viewed as distortion-specific methods by assuming that the prior knowledge of the distortion is provided beforehand or the distortion type is known in advance, which obstructs the application of these methods in the real-world scenario.

In contrast with early distortion-specific methods, recent general-purpose NR-IQA methods can handle images distorted by various distortion types. Based on whether to use subjective scores during the learning process, general-purpose NR-IQA methods can be further categorized as opinion-aware and opinion-free methods. The opinion-free methods require no access to subjective human ratings in the learning process [34,35]. One approach is proposed in [34] to fit a multivariant Gaussian (MVG) model to characterize the regularities of the natural images based on the spatial domain NSS features from natural (distortion-free) images, which is further improved in [35] by combing more perceptual features from image gradient, log-Gabor and color domains.

However, most existing NR-IQA methods fall under the category of opinion-aware methods by learning the computational models based on distorted images and the associated subjective human ratings, which generally takes two stages. The first stage is to extract quality-aware features that represent the image quality from distorted images, while the second stage is to learn a regression model to map the quality-aware features to the final quality score. Intuitively, the choice of the extracted quality-aware features play a vital role in such learning based methods. An ideal set of quality-aware features should be insensitive to image content changes (e.g., illumination change), and sensitive to diverse image distortions. One kind of popular features are based on natural scene statistics (NSS), given the assumption that natural images own regular statistics and various distortion types tend to violate such regularity to certain different degrees. According to such assumptions, many NR-IQA methods take advantage of naturalness of NSS in different domains, such as spatial NSS [36], DCT NSS [37], Wavelet NSS [38], and hybrid of several types of transform domain NSS [39]. Other efficient handcrafted features have also been designed, such as image gradients [40,41], image filter responses [42,43], and image entropies [14,44]. In [40], the authors exploit Laplacian of Gaussian responses and the joint statistics of image gradients as quality-aware features. Apart from first-order derivatives (e.g., image gradients) mentioned above, high-order derivatives of images are also extracted for NR-IQA task in [42,43]. For instance, in [44], the spatial NSS features and free energy principle based features are combined, leading to promising performance in evaluating image quality. With the rapid development of neural network, some opinion-aware methods input raw local image patches for neural networks and attempt to learn quality-aware representation [45,46] instead of designing handcrafted features. For a comprehensive review on NR-IQA methods, the readers can refer to [47].

3. Proposed method

The proposed method is based on the HVS's sensitivity to structural degradation. To make use of the structural information better, we extract a number of first-order and high-order features, which respectively characterize the main structures of images (e.g., edges), and fine-grained details of images (e.g., textures). Such two