Digital Signal Processing 55 (2016) 1–12

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Blind image quality assessment with the histogram sequences of high-order local derivative patterns

Songlin Du, Yaping Yan, Yide Ma *

School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Article history:
Available online 21 April 2016

Keywords:
Blind image quality assessment (BIQA)
Local derivative pattern (LDP)
High-order information
Directional information
Structural degradation

ABSTRACT

Automatic assessment of the perceptual quality of digital image is an important and challenging issue in computer vision. Although human visual system (HVS) is sensitive to degradations on spatial structures, most of the existing methods do not take into account the spatial distribution of local structures. This paper reports a novel approach coined high-order local derivative pattern (LDP) based metric (HOLDPM). In particular, HOLDPM extracts local image structures with LDPs in multi-directions to yield an accurate assessment of image quality. HOLDPM is extensively evaluated on three large-scale public databases. Experimental results demonstrate that HOLDPM is able to achieve high assessment accuracy. Besides, objective assessment result of the HOLDPM is consistent with the subjective assessment result of the HVS. Specifically, the experimental results also indicate that HOLDPM outperforms most of the state-of-the-art methods in distortion specific tests. Additionally, HOLDPM shows competitive overall performance when measured with the weighted average of Spearman rank-order correlation coefficient (SROCC) and the weighted average of Pearson linear correlation coefficient (PLCC) over the test databases.

1. Introduction

With the rapid development and progressive prevalence of the fourth generation (4G) of mobile telecommunications technology, millions of digital images are being transmitted on the communication networks at every moment. Taking photos with a mobile-phone and sharing the photos on social networks have become part of our quotidian existence. However, any operation on image, including acquisition, compression, transmission, and storage, may introduce a variety of distortions. To ensure that the end-users can enjoy a satisfactory perception quality, image quality assessment (IQA) plays a very crucial role in evaluating the performances of image processing methods [1]. IQA aims to describe how image signals are perceived by human observers. The focus is no longer on measurable physical quantities, but rather on how an image is subjectively experienced and whether it is perceived to be of acceptable quality from a subjective point of view [2]. In the last decade, a lot of IQA metrics have been proposed to assess image quality automatically and intelligently [1–3]. According to the degree of the dependency on a reference image, objective IQA metrics can be classified as full reference (FR), no reference (NR), and reduced-reference (RR) methods [3–6]. NR-IQA is also known as blind image quality assessment (BIQA).

Mean squared error (MSE) and peak signal-to-noise ratio (PSNR) are the most widely used FR-IQA metrics, which are achieved by averaging the squared intensity differences between distorted and reference image pixels [7]. The superiorities that MSE and PSNR are simple to calculate and have clear physical meanings facilitated their popularity. However, as an anticipation, good IQA methods should be statistically consistent with the human visual perception, MSE and PSNR don't meet the requirement well. Therefore, to get out of the dilemma, in recent years, the advantages of the known characters of human visual system (HVS) are modeled for image quality assessment. Structural similarity (SSIM) [4] is one of the most approved HVS-oriented IQA metrics. Since SSIM follows the assumption that HVS is highly adapted to extract structural information from the viewing field [4], it can provide a good approximation to perceived image distortion. L. Zhang et al. [8] proposed a feature-similarity (FSIM) index, which employs phase congruency and gradient magnitude to compute a local similarity map for FR-IQA. The similar idea has been used for assessing quality of tone-mapped image [9]. All of MSE [7], PSNR [7], SSIM [4], and FSIM [8] are FR-IQA metrics. Unfortunately, in most applications, information regarding the reference image and the distortion type may be unavailable. So we need to develop algorithms that accurately estimate visual quality without information about the reference image and other prior knowledge.

* Corresponding author.
E-mail address: ydma@lzu.edu.cn (Y. Ma).

http://dx.doi.org/10.1016/j.dsp.2016.04.006

To a certain type of distortion, the degree of distortion can be measured by extracting some specific features. For example, it has been demonstrated that noise can be measured by block homogeneity [10]; blurriness can be measured by cumulative probability [11]; blockiness caused by JPEG (or JPEG2000) compression can be measured by the zero-crossing rate around block boundaries [12]. Actually, these distortion-specific models are not truly “blind”, because they are based on the prior knowledge of distortion types. The aim of this paper is to develop a distortion-generic model which is able to estimate quality of natural image distorted by unknown distortion causes. Studies have shown that high quality natural images exhibit statistical regularities [13,14]: when image structures are distorted or damaged, the corresponding natural scene statistics (NSS) will be altered accordingly. In other words, natural images possess certain regular statistical properties that are altered in the presence of distortion [14,15]. This hypothesis provides a vehicle for BIQA. In recent years, a large plenty of BIQA methods, which employ NSS characteristics to estimate the perceptual quality of natural image, have been proposed. A.K. Moorthy and A.C. Bovik [15] constructed an universal BIQA algorithms coined Distortion Identification-based Image Verity and INtegrity Evaluation (DIIVINE). DIIVINE [15] includes a classification stage and a quality assessment stage. In the classification stage, the type of distortion of an input image is estimated by extracting scene statistics. In particular, the scene statistics are extracted by using the Gaussian scale mixture (GSM) model [16] and the generalized Gaussian distribution (GGD) [17] to simulate the non-Gaussianity (NG) [18] of the wavelet and Gabor coefficients, respectively. In the quality assessment stage, a distortion-specific BIQA algorithm, which uses a classification module, is used to evaluate the distortion-specific quality. A.K. Moorthy and A.C. Bovik [19] also constructed a new statistical model entitled distorted image statistics (DIS), which is an extension of NSS. They demonstrated that each distortion has a unique signature which can be characterized with DIS, and the signature can be used to classify images into distortion categories [19]. Based on DIS, a blind image quality index (BIQI) was proposed by A.K. Moorthy and A.C. Bovik in 2010. M.A. Saad et al. [20] invented what is now known as BLind Image Integrity Notator using DCT Statistics-II (BLIINDS-II) for BIQA. BLIINDS-II also uses GGD to model the non-Gaussian distribution of discrete cosine transform (DCT), and the extracted DCT coefficients are processed by a generalized parametric model. Then the parameters of the model are utilized to estimate image quality scores. BLIINDS-II maps the extracted features to a scalar quality without considering the types of distortions. Instead of employing the most widely used support vector machine (SVM), BLIINDS-II uses a probabilistic graphical model to learn the mapping from extracted features to the scalar quality. Besides, X. Gao et al. [21] constructed an universal blind quality indicators by combining the nonGaussianity (NG), the local dependency (LD), and the exponential decay characteristic (EDC). In addition, they use multiple kernel learning (MKL) to measure the similarity of different features. The NSS features have also been extracted in other transformed domains, such as logarithm domain [22], curvelet domain [23], and contourlet domain [24]. Instead of extracting NSS features in transformed domains, A. Mittal et al. [25] proposed to obtain NSS characters in spatial domain. Their achievement, which is entitled Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE), uses scene statistics of locally normalized luminance coefficients to quantify distortions, thereby leading to a holistic measure of image quality. It is particularly noteworthy that, since local contrast is closely related to the structural information of natural image, W. Xue et al. [26] tried to utilize the joint statistics of the gradient magnitude (GM) map and the Laplacian of Gaussian (LOG) response for BIQA. Specifically, they proposed an adaptive procedure to jointly normalize the GM and LOG features, and proved that the joint statistics of the normalized GM and LOG features have desirable properties in BIQA. By combining the normalized GM and LOG features, they designed a new BIQA metric which is abbreviated as GM-LOG. The GM-LOG model shows highly competitive performance with other state-of-the-art BIQA models. In addition, in our previous work [27], we developed a generalized local ternary pattern (GLTP), which is a novel image structure descriptor, and employed GLTP to measure the degradations of local image structures. Experimental results show that GLTP features can be used to achieve accurate BIQA. Unlike NSS-features-based BIQA models that uses supervised (or semi-supervised) learning algorithms (e.g., SVM [15], graphical model [20], and MKL [21]), P. Ye et al. [28,29] proposed an unsupervised BIQA model by learning a dictionary with raw-image-patches extracted from a set of unlabeled images. Since a soft-assignment coding with max pooling was employed to obtain effective image representations for quality estimation, the proposal was entitled COdebook Representation for No-reference Image Assessment (CORNIA).

Since recent neurological researches on visual perception demonstrate that the HVS is highly adapted to extract structural information for visual content understanding [30,31], we deem that local image structure and its corresponding distribution will be altered in the presence of distortion, and the hypothesis will be verified in Section 2. It is well known that the famous local binary pattern (LBP) can be considered as the binary approximation of the image structure primitives in the early stage of the HVS, and both of the FR-IQA [32] and NR-IQA [33] methods have been proposed by extracting the statistics of structural degradation with LBP. For example, J. Wu et al. [32] employed the traditional uniform rotation invariant LBP [34] to extract structural information to measure the structural degradation on spatial distribution for FR-IQA. M. Zhang et al. [33] proposed a generalized local binary pattern (GLBP) by using a threshold function. The GLBP can effectively represent the statistical distributions of local structures, and the histogram of the GLBP map was used as the quality-aware feature to estimate image quality blindly. Although LBP based metrics [32,33] have made remarkable progresses, they have following inherent drawbacks [35].

1. LBP based metrics actually encode the binary result of the first-order derivative among local neighbors, which are incapable of describing high-order information.
2. LBP is an undirectional operator, so LBP based metrics are incapable of describing direction information.

The fact that LBP is a nondirectional first-order local pattern operator is the primary cause of these drawbacks. In this paper, a new BIQA method coined High-Order Local Derivative Pattern based Metric (HOLDPM) is proposed to address the above problems while inheriting all the strengths of the LBP-employed methods. Our proposal mainly contributes to BIQA in following two aspects.

1. High-order derivative information is encoded as quality-aware features for BIQA.
2. Directional structural information is also encoded as quality-aware features for BIQA.

Taken altogether, the novelty of this work lies in that we propose to use the histogram sequences of high-order local derivative patterns (LDPs) for BIQA. High-order derivative of the structural information and directionality are extracted from raw image as discriminative features. Since our proposal is able to encode more detailed quality-aware features that the first-order local pattern