Neurocomputing 173 (2016) 462–470

Contents lists available at ScienceDirect

Neurocomputing

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

No reference image quality assessment using sparse feature representation in two dimensions spatial correlation ☆

Chuang Zhang [a,b,*], Jing Pan [c], Suting Chen [a], Tingting Wang [a], Dongjiao Sun [c]

[a] Jiangsu Key Laboratory of Meteorological Observation and Information Processing, Nanjing University of Information Science and Technology, Nanjing, China
[b] State Key Laboratory for Novel Software Technology, Nanjing University, China
[c] School of Electronic Engineering, Tianjin University of Technology and Education, China

ARTICLE INFO

Article history:
Received 20 August 2014
Received in revised form
29 December 2014
Accepted 15 January 2015
Available online 2 September 2015

Keywords:
No reference image quality assessment
Feature sparse representation
Spatial correlation
Two dimensions distribution

ABSTRACT

Universal no reference image quality assessment is attracting significant attention in the fields of image processing, visual and machine learning. This study presents a novel method to evaluate the image quality from human subjective scores of the training samples by regressing. The primary characteristic of this novel method is the learning dictionary derived from extracting features of two dimensional spatial correlations of the sample images. Each atom in the dictionary includes 10 elements. They are DMOS (differential mean opinion score), three extracted features and their corresponding image structure patches and PCMs (pixel correlative matrix). The three extracted features are the standard deviation, the gray scale deviation and the distribution width. During the quality assessment, a distorted image is transformed into an image with structural information and partitioned into patches. The patch with the largest feature value is selected and represented sparsely in the learning dictionary. Afterwards the image quality index is obtained by quantifying the sparse representation coefficients, DMOS values and feature values. Comparing with other image quality assessment models, the proposed NSRCIQ method is simple and effective. The resulted IQA scores have not only comparable accuracy, but also high linearity to human perception of image quality. Moreover, the algorithm can be implemented in real-time.

1. Introduction

With the advancement of acquiring and displaying high-resolution images, how to store, share, assess, and cull digital photos faces the significant challenges. Despite the proliferation of captured digital image data, consumers demand better quality image and video acquisition and display. It is highly desirable to be able to automatically and accurately predict visual signal quality. Quality assessment algorithms can be used to improve picture quality via quality-aware processing, computing and networking. Especially, no reference image quality assessment (NRIQA) plays a crucial role in image processing and analysis, while the standard pictures cannot be acquired. Scholars have developed numerous universal and robust algorithms to evaluate the image quality without reference images in recent years [1–3]. The feature extraction and mapping model are two primary components in estimating image quality without the undistorted image [4]. There are two classes of approaches, direct mapping with extracted features and subsequent mapping with distortion classified, for NRIQA metrics. Direct mapping method is called universal metric, which performs without the information about the specific distortion. As this universal metric with a wide applicability, the scientific community has developed some algorithms [5–9]. Although the most efficient and accurate metric has not been found out, several approaches for NRIQA perform consistently with the subjective perception across various kinds of distortions [10–12].

It is natural scene statistics (NSS) [13] that has been widely used in universal blind quality assessment metrics [14–17]. NSS, which describes particular structures and properties in images, is used to present the natural scene images [18]. For distorted images that can change NSS, NSS is used as feature extraction in NRIQA

☆The present study is supported by the Priority Academic Program Development of Jiangsu Higher Education Institutions project, Jiangsu Key Laboratory of Meteorological Observation and Information Processing open project (kdxs1203), State Key Laboratory for Novel Software Technology (Nanjing University) Project (KFKT2012B20), and the National Natural Science Foundation of China (Grant no. 61271412).

* Corresponding author at: Jiangsu Key Laboratory of Meteorological Observation and Information Processing, Nanjing University of Information Science and Technology, Nanjing, China.
E-mail address: zhch_76@163.com (C. Zhang).

http://dx.doi.org/10.1016/j.neucom.2015.01.105
0925-2312/

metric. So the mapping model focuses on estimating the distorted extent by calculating the deviation between the changed NSS and the NSS obtained from undistorted natural images [19]. NSS is obtained by a scale-space-orientation decomposition (such as wavelet transform) to images. Although these NSS in the wavelet domain can provide the significant structural information, they cannot find out the image block with the most prominent structural change by using wavelet transform to a whole image. NSS neglects the influence of the illumination on the structural information, while the structures of the objects in the scene are independent of the illumination. Intuitively, one interesting question is how to develop an effective and efficient NRIQA algorithm that only uses the most prominent structure information.

Consequently, a NRIQA metric is proposed that only uses the most prominent structure information. As the luminance of the surface of an object being observed is the product of the illumination and the reflectance in an image, the structures of the objects in the scene are independent of the illumination. For separating the influence of the illumination and avoiding the luminance and contrast vary across a scene, the local average luminance and contrast are used to explore the structural information in an image [20]. Subsequently, each structure image patch, which is independent of the illumination, is calculated to obtain pixel correlative matrix (PCM) [21]. At the same time, the image patch is selected by finding out the maximum PCM standard deviation. Here, the maximum PCM standard deviation means the most prominent structure information. Then, a dictionary is constructed by combining a set of extracted features from the image patches, which have the most abundant structure information in their respective images with different types of distortions. Finally, the extracted features of an evaluated image are represented in the dictionary above and NRIQA score is applied by mapping the correlative coefficients to the corresponding subject quality (differential mean opinion score, DMOS).

The rest of the paper is organized as follows. The feature extraction by PCM is described in Section 2 in detail. Then the proposed NRIQA algorithm using sparse representation is in Section 3. Experiments and discussions are detailed in Section 4. Finally, Section 5 concludes the paper.

2. Pixel correlative matrix and feature extraction

It is well known that the human visual system (HVS) is highly adapted to extract structural information from the viewing field. It follows that a measure of structural information change can provide a good approximation to perceived image distortion [22]. Experiments show that image degradations as perceived changes in structural information. Simultaneously, the HVS is sensitive to the relative luminance change, and not the absolute luminance change. To obtain the structural information, which is independent of illumination, the local average luminance and contrast are used to explore the structural information in an image according to the SSIM method [23].

2.1. Eliminating the influence of the illumination

Suppose $\boldsymbol{X}$ is a nonnegative image signal, which is an arbitrary spatial patches extracted from each image. The task of eliminating the influence of the illumination is separated into three parts as follows. First, the luminance of each patch is computed. Assuming discrete signals, this is estimated as the mean intensity:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

where N denotes the number of pixels in the image patch.

Second, we remove the mean intensity from the signal. And the resulting signal is $\boldsymbol{X}-\mu_x$.

The standard deviation (the square root of variance) is used as an estimate of the signal contrast. The standard deviation in discrete form is given by

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu_x)^2} \tag{2}$$

Third, the signal is normalized (divided) by its own standard deviation. And the image patch structural information can be expressed as

$$\boldsymbol{X}' = (\boldsymbol{X}-\mu_x)/\sigma_x \tag{3}$$

An image (from IVC database [23]) and its own structural information are shown in Fig. 1.

2.2. Pixel correlative matrix

The PCM can be obtained by the following steps [21]:

(1) Build a matrix with a size of 256×256;
(2) The grayscale of pixel (i,j) is $g1$ and the grayscale of pixel $(i,j+l)$ is $g2$, the point $(g1, g2)$ in the matrix is set 1, where l is the distance between the two pixels.

Fig. 1. An image and its structural information.