Information Sciences 273 (2014) 196–211

Contents lists available at ScienceDirect

Information Sciences

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

Image quality assessment based on matching pursuit

Richang Hong [a], Jianxin Pan [b], Shijie Hao [a], Meng Wang [a,*], Feng Xue [a], Xindong Wu [a]

[a] School of Computer and Information, Hefei University of Technology, 230009 Hefei, China
[b] Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, 518055 Shenzhen, China

ARTICLE INFO

Article history:
Received 25 April 2012
Received in revised form 19 February 2014
Accepted 3 March 2014
Available online 21 March 2014

Keywords:
Objective image quality assessment
Matching pursuit
Structural information
2D Gabor dictionary

ABSTRACT

The objective assessment of image quality is an essential part of many visual processing systems. The challenge lies in evaluating the image quality consistently under subjective perceptions. In this paper, we propose a novel image quality metric based on the matching pursuit algorithm. Under the principle of structural information distortion, we assume that various structure data contributes differently to the single image quality score. Specifically, we decompose the reference image using matching pursuit with a separable 2D Gabor dictionary, thus obtaining structural information, and develop a characterization for this information and its importance. We then discuss the relationship between the structural distortion intensity and the subjective quality measurement on the energy scale. In experiments, we compare the performance of our algorithm with both subjective ratings and state-of-the-art objective methods on image datasets with multi-type distortions. The experimental results validate our proposed method.

1. Introduction

Visual quality may be degraded during image and video procedures such as acquisition, processing, coding, storage, and transmission. The purpose of image quality assessment (QA) is to develop strategies and algorithms to accurately evaluate the quality and make this consistent with our subjective perception. However, although this is pivotal in many visual systems, such as image retrieval and surveillance systems, it remains a challenging and unresolved issue. In general, we can classify the metrics for image QA as *subjective* or *objective* [2]. As we all know, human beings are the best assessors of image quality. The mean option score (MOS) is one kind of subjective assessment that is widely used in image and video QA. However, in practice, MOS is time-consuming, labor-intensive, and environment-limited. Besides these limitations, the subjective assessment may be affected by various factors, e.g., the mood of the assessors and the experimental setting. Therefore, an objective image quality metric that can automatically and precisely evaluate the image quality is highly desirable [1].

Objective image quality metrics can be classified as *full-reference*, *no-reference*, or *reduced-reference* based on the original image used (we take it for granted that the original image is ‘perfect’ or of ‘high quality’ and define it as the reference): (1) *full-reference* means that the original image is fully known; (2) *no-reference* (also called *blind*) means the reference image cannot be obtained; (3) *reduced-reference* means the reference image is only partially known (in other words, some key features of the image are given). In this study, we focus on full-reference image QA. Recently, many algorithms have been

* Corresponding author. Tel.: +86 55162904883 (M. Wang).
E-mail addresses: hongrc.hfut@gmail.com (R. Hong), hsj.hfut@gmail.com (S. Hao), eric.mengwang@gmail.com (M. Wang), feng.xue@hfut.edu.cn (F. Xue), xwu@hfut.edu.cn (X. Wu).

http://dx.doi.org/10.1016/j.ins.2014.03.009
0020-0255/

presented that tackle objective image QA in the presence of references, i.e., under the paradigm of *full-reference*. We categorize these into three classes:

1. Metrics based on **pixel statistics**, which use various statistics to represent the quality of the image, such as the mean square error, root mean square error (RMSE), signal-to-noise ratio, and peak signal-to-noise ratio (PSNR). These metrics cannot completely reveal the characteristics of human vision or agree with the feelings of humans [1,3,4,30,31]. Although they have some systematic drawbacks, these metrics are still widely used in many situations that have low requirements, as they are relatively simple to perform.
   (a) Metrics based on the **human visual system (HVS)**. The application of the characteristics of human vision to image QA was proposed by Mannos and Sakrison [5] in 1974. They developed the contrast sensitivity closed formula (based on psychophysical measurements), which is still used in HVS models. Karunasekera and Kingsbury have researched different types of distortion models separately, and weighted-summed the distortion errors based on the different sensitivities of human vision to obtain their final results [6]. Chou and Li proposed a metric based on just-noticeable distortion (JND), and decomposed images into multiple channels to evaluate a grayscale image's JND/MND (minimally noticeable distortion) distribution. They made full use of the background luminance conceal and grain conceal characteristics of the human visual system, and then designed a sub-zone distributing arithmetic to remove the visual redundancy calculated by the JND/MND model to get a visually lossless or near-lossless affection [7]. Watson reported a visual error assessment metric based on the discrete cosine transformation (DCT), in which the coefficient of each DCT quantitative error is magnified or lessened with the visual sensitivity [8]. Mayache et al. compared the three main image QA metrics based on HVS [9]. Although the psychophysical measurement of HVS is mostly accepted in the community, the complexity of HVS and the finite cognizance of human beings prevents this metric from further development [1,10,34].
2. Metrics based on **structural distortion of images**. This idea was developed by Wang et al. [1]. On the assumption that "human visual perception is highly adaptive for extracting structural information from a scene", they proposed the Structural Similarity (SSIM) metric, which compares the structural similarity of the original image and the distorted one. They regard "structural information" as an object attribute that is independent of the luminance and contrast of the image. The arithmetic of Mean Structural Similarity (MSSIM) divides the image information into three different aspects: luminance, contrast, and structure, which are calculated individually. In this context, an image is divided into small blocks, and the distortion of each block is computed using the luminance, contrast, and structure information, and the final quality score of the image is the mean value of the blocks [1]. Shnayderman et al. proposed assessing images based on the singular value decomposition (SVD) of the matrix of the image, and obtained good results [11]. Other new measures have been reported recently [10,12–15].

Generally, an image may be segmented into structural and texture components, with the structural information playing a more important role in many cases. Natural images usually contain abundant structural elements that carry important information about the objects' structure for the visual scene [1]. It is easy to observe that an image contains structural information under different scales, which generally comfort to the pattern of human vision. Therefore we argue that both the characterizations of structural information and their importance should be taken into consideration in the process of image QA. To this end, we try to decompose the original image to gain the structure information in the order of importance, and give the characteristics of importance for different structure.

In this paper, we explore the feasibility of developing a new image QA metric using the matching pursuit (MP) technique. Under the assumptions above, we develop a characterization for the structural information of an image, and characterize its importance by projecting the image onto an over-complete separable 2D Gabor basis set using MP. We then discuss the relationship between the structural distortion magnitude and the subjective quality magnitude. The remainder of this paper is organized as follows: Section 2 introduces the metric of MP, the dictionary used in our approach, and the modifications made to MP for image QA. In Section 3, we develop a new metric for the local and global structural distortion measurement, and describe the final measure: an image quality score based on MP (MP_Q). In Section 4, we compare our experimental results with some other metrics and analyze their performance. Section 5 concludes the paper.

2. Matching pursuit for our image quality measure

Matching pursuit, which decomposes the signal into optimized bases selected from a redundant dictionary, is a general procedure for computing the signal representations and extracting important structural information from the signal. The bases are chosen to best match the signal structure and extract relevant structural information.

2.1. Matching pursuit

The MP algorithm was developed by Mallat and Zhang [16]. This algorithm takes advantage of the projection pursuit as a measure for time–frequency analysis. Simply speaking, its kernel arithmetic is that a signal is projected on an optimized basis, and the most important part of the signal, under the rule of the largest energy, is chosen.