ARTICLE IN PRESS

Signal Processing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Signal Processing

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

Image quality assessment based on the space similarity decomposition model ☆

Yang Yang *, Jun Ming

Key IC&SP Lab of Ministry of Education, Anhui University, Hefei 230039, China

ARTICLE INFO

Article history:
Received 8 August 2014
Received in revised form
9 March 2015
Accepted 23 March 2015

Keywords:
Image quality assessment
Vector decomposition
Structural similarity
Weber–Fechner law

ABSTRACT

In the image quality assessment (IQA) research field, the structural similarity index measurement (SSIM) method and human visual system (HVS) model have received much attention. However, this paper shows that the definition of the luminance comparison function in SSIM conflicts with the Weber–Fechner law in HVS. To reconcile this contradiction, we propose a space similarity decomposition model based on the Weber–Fechner law and SSIM's structure parameter to decompose the image into mean value, similarity, and vertical components. Experimental results show that the vertical component is the main factor of subjective perception and should be used by itself to assess image quality. This method not only can be used to independently assess image quality, it can also be combined with other IQA methods as a preprocessing step.

1. Introduction

With the exponential growth of image content available online, assessing the quality of images has become an increasingly important task [1–5]. Recently, web image retrieval has attracted much attention and achieved great success in many applications such as multimedia question–answer research [6–11]. For all images retrieved from the web, we would like to automatically judge their quality and select those with good visual quality. Hence, the image quality assessment (IQA) method can be applied to automatically and efficiently judge image quality. Presently, the main research direction in this field is objective IQA, which aims to accurately reflect the subjective quality of human vision and focuses on reducing the deviation between subjective and objective quality assessment results [1–3].

In full reference IQA research field, the human visual system (HVS) [12–14] and structural similarity index measurement (SSIM) methods [15–17] have received extensive recognition. In general, the HVS model consists of three sub models: the luminance visual nonlinear perception model that is derived from the Weber–Fechner experiment, the multi-channel decomposition response model that is derived from the visual resolution experiment, and the visual masking model that is derived from the multiple differences of luminance visual perception experiment. The luminance visual nonlinear perception model, also called the nonlinear perception model, reveals the logarithm nonlinear relationship between subjective luminance perception and objective luminance value. However, it does not consider the structural similarity information of an image.

The SSIM method defines three comparison functions (structural, luminance, and contrast comparison), where the structural comparison function measures the structural similarity information of between two images and its essence is the cosine value of the angle between two images

☆ This work was supported in part by the National Nature Science Foundation of China (No. 11374289, No. 60872162), by the Young Research Foundation of Anhui University under Grant (KJQN1012) and by the Doctoral Scientific Research Foundation (J01001319).

* Corresponding author. Tel.: +86 18055169268.
E-mail addresses: skyyang@mail.ustc.edu.cn (Y. Yang), video@ahu.edu.cn (J. Ming).

http://dx.doi.org/10.1016/j.sigpro.2015.03.019
0165-1684/© 2015 Elsevier B.V. All rights reserved.

that removed mean value separately. In fact, the structural similarity idea has emerged in other research areas, such as image categorization and recognition research [18–22]. One key application is model matching, which often needs to calculate the structure similarity between a constructed model and the extracted information of an image. One useful way to do this is to calculate the structural similarity information, which is the same as the structural comparison function in SSIM. However, the definition of the luminance comparison function in SSIM not reflects the subjective luminance perception and objective luminance value's logarithm nonlinear relationship that has been demonstrated by the Weber–Fechner law.

Hence, motivated by the structural similarity idea in SSIM and nonlinear perception model in the Weber–Fechner law, we propose a space vector decomposition model that decomposes the image into mean value, similarity, and vertical components. The experiments demonstrate that the error image mean μ_e and error similarity $\tilde{\mathbf{s}}_{e=}$ are similar to the original image background component μ_0 and original image alternating value $\tilde{\mathbf{s}}_0$, respectively, hence they don't bring the new subjective perception and should be extracted and removed before assessing image quality. However, the subjective perception of vertical error $\tilde{\mathbf{s}}_{e\perp}$ is different from that of the original image, hence it is the key factor that influences subjective perception, and should be extracted and used to assess image quality. The proposed method could also be used as a preprocessing step in combination with other IQA schemes. Furthermore, the proposed method could be used to automatically judge the quality of images from the Web in image retrieval [6–11].

This paper is organized as follows. Section 2 analyzes the HVS model and SSIM scheme. Section 3 elaborates on the space similarity decomposition model. The performance of the proposed method is evaluated and compared with other methods in Section 4, and the conclusion is presented in Section 5.

2. HVS model and SSIM method analysis

Currently, two methods have received extensive recognition in IQA research; one is the HVS-based model and the other is the SSIM-based model.

Because objective IQA aims to be consistent with subjective visual perception, the HVS model is based on the human visual system. This model consists of the luminance visual nonlinear perception, visual multi-channel decomposition, and visual masking models. The luminance visual nonlinear perception model reflects the logarithm nonlinear relationship between the subjective luminance perception and objective luminance value that is derived from the Weber–Fechner law. This model is written as

$$\mathbf{h} = f(\mathbf{s}) = k \lg \mathbf{s} + k_0, \tag{1}$$

where $\mathbf{s}$ is objective luminance, $\mathbf{h}$ is subjective visual perception luminance, k and k_0 are the constant.

The visual multi-channel decomposition model demonstrates the effect of frequency domain information on subjective visual perception that is derived from visual resolution experiments. One typical representation is the contrast sensitivity function,

$$A(f) = 2.6(0.0192 + 0.114f)\exp\left[-(0.114f)^{1.1}\right], \tag{2}$$

where f is the center frequency (cycle/degree) and $A(f)$ is the degree of subjective sensitivity.

The visual masking model demonstrates the effect of a variety of luminance difference interactions on subjective visual perception,

$$\mathbf{m}_i = g(\mathbf{s}_i)\phi(\mathbf{s}_i), \tag{3}$$

where

$$\phi(\mathbf{s}_i) = \begin{cases} 1 & \text{if} \mathbf{s}_i \geq T_i \\ 0 & \text{else} \end{cases},$$

and T_i is the visual perception threshold.

However, as the HVS model has been further explored, its algorithms have become more complex and the experimental results less adaptable. In addition, the HVS model does not exploit the structural information of the image.

The SSIM model considers the structural characteristic of the image and consists of the structure, luminance, and contrast comparison functions

$$s(x,y) = \frac{\sigma_{xy}}{\sigma_x\sigma_y}, \quad l(x,y) = \frac{2\mu_x\mu_y}{\mu_x^2+\mu_y^2}, \text{ and } c(x,y) = \frac{2\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2},$$

respectively. Here, μ_x and μ_y are the mean luminances of the original image $\mathbf{x}$ and distortion image $\mathbf{y}$, σ_x^2 and σ_y^2 are the variance of $\mathbf{x}$ and $\mathbf{y}$, and σ_{xy}^2 is the covariance of $\mathbf{x}$ and $\mathbf{y}$. Finally, the IQA parameter is $S(x,y) = l(x,y)^{\alpha} * c(x,y)^{\beta} * s(x,y)^{\gamma}$ $S(x,y) = l(x,y)^{\alpha} * c(x,y)^{\beta} * s(x,y)^{\gamma}$, in which α, β, γ are the positive value and they used to adjust the weighting of three functions separately. The SSIM structural comparison function reflects the structural information of the image, which is also successfully used in model matching algorithms [18–22]. Here, the essence of the structural comparison function in SSIM is the cosine value of the angle between two images that removed mean value separately. However, the definition of the SSIM luminance comparison function conflicts with the logarithm nonlinear relationship of the subjective luminance perception and objective luminance value in the Weber–Fechner law.

3. Space similarity decomposition model

According to the Weber–Fechner law [23], there is a logarithm nonlinear relationship between subjective visual perception luminance $\mathbf{h}$ and image objective luminance $\mathbf{s}$, namely $\mathbf{h} = f(\mathbf{s})$. Given the existing observation environment calibration in subjective IQA, we define μ_b be the observed environment luminance, and the subjective luminance perception increment is then $\Delta\mathbf{h} = f(\mu_b + \mathbf{s}) - f(\mu_b)$. This can be approximated as $\Delta\mathbf{h} = f'(\mu_b)\mathbf{s}$ using a Taylor series approximation. For original image $\mathbf{s}_0$ and distortion image $\mathbf{s}_1$, the difference between the two subjective luminance perception increments is $\Delta\mathbf{h}_0 - \Delta\mathbf{h}_1 = f'(\mu_b)(\mathbf{s}_0 - \mathbf{s}_1)$. Defining the error image as $\mathbf{s}_e = \mathbf{s}_0 - \mathbf{s}_1$, we have

$$\Delta\mathbf{h}_0 - \Delta\mathbf{h}_1 = f'(\mu_b)\mathbf{s}_e. \tag{4}$$

Therefore, Eq. (4) states that the difference between the subjective luminance perceptions in the two images is proportional to their objective error image $\mathbf{s}_e$. Hence, $\mathbf{s}_e$ is