J. Vis. Commun. Image R. 34 (2016) 230–235

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Short Communication

Perceptual similarity between color images using fuzzy metrics ☆

Svetlana Grečova [a,*], Samuel Morillas [b]

[a] Department of Mathematics, University of Latvia, Latvia
[b] Instituto Universitario de Matemática Pura y Aplicada, Universidad Politécnica de Valencia, Spain

ARTICLE INFO

Article history:
Received 18 February 2014
Accepted 5 April 2015
Available online 11 April 2015

Keywords:
Color imaging
Fuzzy logic
Fuzzy metrics
Perceptual image similarity
Color similarity
Perceptual observations
Low level image processing
Color image quality

ABSTRACT

In many applications of the computer vision field measuring the similarity between (color) images is of paramount importance. However, the commonly used pixelwise similarity measures such as Mean Absolute Error, Peak Signal to Noise Ratio, Mean Squared Error or Normalized Color Difference do not match well with perceptual similarity. Recently, it has been proposed a method for gray-scale image similarity that correlates quite well with the perceptual similarity and it has been extended to color images. In this paper we use the basic ideas in this recent work to propose an alternative method based on fuzzy metrics for perceptual color image similarity. Experimental results employing a survey of observations show that the global performance of our proposal is competitive with best state of the art methods and that it shows some advantages in performance for images with low correlation among some image channels.

1. Introduction

Many applications in the fields of image processing and computer vision use image similarity measures for different purposes [1]. In some cases the objective is the very measurement of the similarity itself globally or partially in the images, but other times the similarity is used to assess the performance of an image processing method. For instance, in image filtering, the common process to measure the performance of a filtering method is the following: an original image is corrupted artificially with noise, then it is filtered with the method under study and it is measured how similar is the filtered image to the original one. This allows to properly adjust filter parameters for optimal performance, to assess different filter configurations as well as to compare the performance of different filtering methods. An analogous approach is used in other image processing procedures such as image compression, image demosaicing or video de-interlacing. Therefore, the similarity measure used highly influences the whole process.

The most common similarity measures used in this context are based on a pixelwise approach, such as the Mean Absolute Error (MAE), the Mean Squared Error (MSE), the Peak Signal to Noise Ratio (PSNR) or the Normalized Color Difference (NCD) (which is the MSE in the Lab color space). However, these measures do not match well with perceptual observations and, as the MSE, some of them have other concerns [2].

During the last twenty years series of works have addressed the problem of defining image similarity measures that match human perceptual similarity. First works in this issue include the Weighted Signal to Noise Ratio (WSNR) [3] which simulates the human visual system properties by filtering both the reference and distorted images with contrast sensitivity functions and then compute the SNR. Other measures [4,5] assess shifts in image luminance, differences in the frequency domain and changes in edges. Instead of luminance, some metrics [6–8] specifically target color in images. Other metrics [9,10] embed a hidden signal in an image, introduce an impairment and measure its quality. Besides, to detect similarity between images their histograms have been used [11,12].

More recently, in [13,14] a similarity measure for gray-scale images that matches well with perceptual similarity has been introduced (UQI-Universal Quality Index and SSIM-Single-scale Structural Similarity Index). This method could be applied in color images in a componentwise fashion, that is, independently in each color channel and then averaged. However, it is well-known that the correlation among the color image channels should be taken into account and this approach cannot provide optimal performance [1], as we show in this paper. This similarity measure is extended to the Multiscale Structural Similarity Index (MSSIM) in [15]. In turn, in [16], a color comparison criterion is combined with

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author at: Reznas 10 k-1, Riga LV-1019, Latvia.
E-mail addresses: grecova.svetlana@gmail.com (S. Grečova), smorillas@mat.upv.es (S. Morillas).

http://dx.doi.org/10.1016/j.jvcir.2015.04.003
1047-3203/

MSSIM. In the approach [17], SSIM scores are weighted by region type. And, in [18], a two staged wavelet based Visual Signal to Noise Ratio (VSNR) was defined based on the low-level and the mid-level properties of human vision.

In this paper, we introduce a method for color image similarity that matches perceptual similarity. Our method follows a procedure inspired in [13,14] as follows: the images are processed with sliding patches so that a number of small image portions are compared and the similarity between two images is obtained by averaging the similarities of all portions. In each pair of patches three different factors are compared separately and then combined: contrast, structure and luminance. The particular expressions used in [13,14] for these three factors cannot be directly generalized from gray-scale images to color images, so we propose our own expressions to measure them. Experimental results employing perceptual similarity observations show that our approach is able to outperform classical similarity measures, is competitive with best state-of-the-art methods, and shows some advantages in performance for images with low correlation among some image channels.

In the following section we detail the proposed method. Section 3 contains the experimental results and discussion. Finally, Section 4 presents the conclusions.

2. Proposed image similarity measure

Let $\mathbf{X}$ denote a RGB image and W be the sliding patch of finite size $q \times q = n$ used to process the image. The image pixels in W, $\mathbf{X}_W$, are denoted as $\mathbf{x}_i(l),\ i = 1, \ldots, n$ where $l = 1,\ 2,\ 3$ denotes the R, G, and B channels, respectively. Notice that $\mathbf{x}_i$ can be processed as a three component vector.

We measure the similarity between images $\mathbf{X}$ and $\mathbf{Y}$ as the average of the similarities of the image patches $\mathbf{X}_W$ and $\mathbf{Y}_W$ obtained when sliding the patch along every image row. To measure the similarity between two patches in the same image location we measure three different similarities: contrast, structure and luminance. In so doing, we need to measure the similarities between all image color pixels $\mathbf{x}_i$ and $\mathbf{y}_i$ in $\mathbf{X}_W$ and $\mathbf{Y}_W$, respectively, and the mean color vector in each patch, $\overline{\mathbf{x}_W}$ and $\overline{\mathbf{y}_W}$. We denote these similarities by $M_{\mathbf{x}_i}$ and $M_{\mathbf{y}_i}$ and we measure them by employing the fuzzy metric used in [19–22] for its high sensitivity to edges as follows:

$$M_{\mathbf{x}_i} = M(\mathbf{x}_i, \overline{\mathbf{x}_W}, t) = \prod_{l=1}^{3} \frac{\min(x_i(l), \overline{\mathbf{x}_W}(l)) + t}{\max(x_i(l), \overline{\mathbf{x}_W}(l)) + t}, \quad i = 1, \ldots, n, \tag{1}$$

where $t > 0$ and

$$\overline{\mathbf{x}_W} = \frac{1}{n}\sum_{j=1}^{n} x_j, \quad l = 1, 2, 3 \tag{2}$$

Through an analogous computation in the image $\mathbf{Y}$ we obtained the similarities $M_{\mathbf{y}_i},\ i = 1, \ldots, n$. Notice that $M_{\mathbf{x}_i}$ and $M_{\mathbf{y}_i}$ are fuzzy similarities that take value in $[0, 1]$.

2.1. Contrast

Contrast can be seen as the largest difference observed in $\mathbf{X}_W$ and $\mathbf{Y}_W$. We can measure contrast in $\mathbf{X}_W$ using $M_{\mathbf{x}_i}$ as $C_{\mathbf{X}_W} = \max(M_{\mathbf{x}_i}) - \min(M_{\mathbf{x}_i}),\ i = 1, \ldots, n$, and analogously for $\mathbf{Y}_W$. Then, the fuzzy similarity between the contrasts is given by

$$SC(\mathbf{X}_W, \mathbf{Y}_W) = 1 - \mid C_{\mathbf{X}_W} - C_{\mathbf{Y}_W} \mid . \tag{3}$$

2.2. Structure

Structure describes how the differences between the pixels in a patch are distributed spatially. Therefore, for this aspect we average the fuzzy similarities of $M_{\mathbf{x}_i}$ and $M_{\mathbf{y}_i}$ as follows.

$$SS(\mathbf{X}_W, \mathbf{Y}_W) = \frac{\sum_{i=1}^{n} 1 - \mid M_{\mathbf{x}_i} - M_{\mathbf{y}_i} \mid}{n}. \tag{4}$$

2.3. Luminance

To compare image luminance we propose to use spherical coordinates computed from RGB values [23]. Luminance correspond with the radius parameter given by

$$L\mathbf{x}_i = \sqrt{\mathbf{x}_i(1)^2 + \mathbf{x}_i(2)^2 + \mathbf{x}_i(3)^2} \tag{5}$$

The luminance similarity between $\mathbf{X}_W$ and $\mathbf{Y}_W$ is obtained through the corresponding expression in [13] as

$$SL(\mathbf{X}_W, \mathbf{Y}_W) = \frac{2\overline{L_{\mathbf{X}_W}}\,\overline{L_{\mathbf{Y}_W}}}{\overline{L_{\mathbf{X}_W}}^2 + \overline{L_{\mathbf{Y}_W}}^2} \tag{6}$$

where $\overline{L_{\mathbf{X}_W}}$ and $\overline{L_{\mathbf{Y}_W}}$ are the mean luminance in each patch. In the case that $\overline{L_{\mathbf{X}_W}} = \overline{L_{\mathbf{Y}_W}} = 0$ we assign $SL(\mathbf{X}_W, \mathbf{Y}_W) = 1$.

Finally, the similarity between $\mathbf{X}_W$ and $\mathbf{Y}_W$ results from combining the three previous measures as follows

$$S(\mathbf{X}_W, \mathbf{Y}_W) = SC(\mathbf{X}_W, \mathbf{Y}_W)^{\alpha} \cdot SS(\mathbf{X}_W, \mathbf{Y}_W)^{\beta} \cdot SL(\mathbf{X}_W, \mathbf{Y}_W)^{\gamma} \tag{7}$$

where $\alpha,\ \beta,\ \gamma > 0$ are parameters used to adjust relative importance of three components. As commented above, the average of all $S(\mathbf{X}_W, \mathbf{Y}_W)$ provides the similarity between $\mathbf{X}$ and $\mathbf{Y}$, that will be high only if the three similarities are high.

Finally, we would like to point out that in each processing patch the number of operations is proportional to the number of pixels, so for the whole method we have also a linear computational cost.

3. Experimental study

In order to study the performance of our proposal and also to compare with other approaches we make a comparison with respect to a survey of perceptual observations as follows.

We have chosen the four color bmp images in Fig. 1: Goldhill, Lenna, Baboon, and Parrots. To better appreciate low resolution differences we have taken a small part of 68×68 pixels of the original images. We have applied a series of 10 different distortions to each of the test images. The distortions applied over the image Parrots along with the software use in each case, which are shown in Fig. 2, are the following.

1. jpg compression of ratio 20% (MS Picture Manager).
2. Increase brightness by 15% (MS Picture Manager).
3. Increase contrast by 15% (MS Picture Manager).
4. Gaussian blur with radius 1.5 (Corel Draw X5).
5. Addition of 5% of impulsive noise (imnoise function from Matlab).
6. Addition of white Gaussian noise with standard deviation equals to 10% of the maximum value in the channels (imnoise function from Matlab).
7. Filtering of original image with [24].
8. Addition of Gaussian noise as in (6) and filtering with [24].
9. Filtering of original image with Vector Median Filter (VMF) [25].
10. Addition of 5% of impulsive noise as in (5) and filtering with Vector Median Filter (VMF) [25].

In the survey, we asked independent observers to rank the 10 distorted images with respect to its similarity to the original image (1st the most similar, 10th the least). We did this through a questionnaire available on the internet address [27] to get as many answers as possible. We received 108 complete answers. We processed them to remove outliers using boxplot and we found 4