Available online at www.sciencedirect.com
SCIENCE DIRECT®

Pattern Recognition Letters 27 (2006) 1054–1066

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Content-based image enhancement in the compressed domain based on multi-scale α-rooting algorithm

Sangkeun Lee *

Digital Media Solutions Lab, Samsung Information Systems America, 3345 Michelson Dr. # 250, Irvine, CA 92612, USA

Received 20 February 2005; received in revised form 28 November 2005
Available online 20 February 2006

Communicated by G. Sanniti di Baja

Abstract

A simple multi-scale image enhancement algorithm for compressing image dynamics and enhancing image details in the discrete cosine transform (DCT) domain is presented. First, an image is separated into illumination and reflectance components. Next, the illumination component is manipulated adaptively for image dynamics by using a content measure. The content measure using the energy distribution of the DCT coefficients is defined directly in each DCT block of an image. Then, the reflectance component is altered by a multi-scale α-rooting method for enhancing image details based on human visual perception. The main advantage of the proposed algorithm enhances the details in the dark and the bright areas with low computations without boosting noise information and affecting the compressibility of the original image since it performs on the images in the compressed domain. In order to evaluate the proposed scheme, several base-line approaches are described and compared using enhancement quality measures.

PACS: 07.05.Pj; 07.05.Rm; 42.30.Tz; 42.30.Va

Keywords: Image dynamics; Illumination; Reflectance; Discrete cosine transform (DCT); α-Rooting; Multi-scale structure

1. Introduction

Image enhancement involves processing an image to make it satisfactory to human viewers. Image enhancement is desirable in a number of contexts, one of these contexts is image enhanced by modifying its dynamic range or contrast or both.

A scene often contains a large dynamic range that cannot be adequately captured by the imaging devices. In the dynamic range compression process, we apply image processing techniques to recover the effects of the full dynamic range from the captured image. After the dynamic range compression is achieved, however, some details may become clustered together within a small intensity range, especially in very dark or bright regions of the image. A contrast/detail enhancement process adjusts the local contrast in different regions of the image so that the details in dark or bright regions are again brought out and revealed to the human viewers. Therefore, dynamic range compression and contrast enhancement together improve the quality of the image. The enhanced image looks closer to real natural scenes, clearer with more details, and more visually pleasing.

Many techniques have been developed for dynamic range compression and contrast enhancement (Gonzalez and Wintz, 1987; Gordon, 1989; Jain, 1989; Ji et al., 1994; Jobson et al., 1997). According to the survey of available techniques found in (Gonzalez and Wintz, 1987; Jain, 1989; Aghagolzadeh and Ersoy, 1992), the existing techniques can be categorized into two classes (Agaian et al., 2001): spatial-domain processing and compressed-domain

* Tel.: +1 949 975 7266; fax: +1 949 975 7299.
E-mail address: sangkny@yahoo.com

0167-8655/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2005.12.004

processing. The former method operates directly on pixels, and most of those methods are based on gray-level histogram modifications (Jain, 1989; Kim, 1997; Chen and Ramli, 2003), while others are based on local contrast measure and edge information (Ji et al., 1994; Beghcladi and Negrate, 1989). In particular, the *Retinex* approach has been introduced and successfully applied to image dynamic range compression (Land and McCann, 1971; Jobson et al., 1997; Funt et al., 2000). The compressed domain methods operate directly on transforms of the image, such as the Fourier, wavelet, and cosine transforms, by analyzing and modifying their spectral coefficients of the transforms. For example, α-rooting (Peli, 1990; Kogan et al., 1998) (which is called magnitude reduction method), modified unsharp masking (Jain, 1989; Pratt, 1991), and multi-scale contrast measure (Tang et al., 2003) algorithms are belonged to the compressed domain methods. The advantages of frequency domain processing techniques are (1) low complexity of computations and (2) easy to view and manipulate the frequency composition of the image, without direct reliance on spatial information (Agaian et al., 2001). However, it is reported that the frequency based methods introduce objective blocking artifacts, and cannot simultaneously enhance all parts of the image very well (Jain, 1989; Aghagolzadeh and Ersoy, 1992; Gonzalez and Wintz, 1987).

In this paper we provide a simple algorithm to compress image dynamic range and to enhance image details in the DCT compressed domain. First, for compressing dynamics of an image, the proposed method employs a basic concept of the Retinex theory, and applies it only to DC components of the DCT coefficients because the DC components represent local image brightness corresponding to the illumination component of Retinex theory. Then, for improving image details, we define spectral bands according to spatial frequency property of each block of the input image, split the bands into several regions, and manipulate the DCT coefficients of each region differently based on α-rooting algorithm with a form of a multi-scale structure. It is important to note that a multi-scale enhancement scheme may be found in (Tang et al., 2003; Kogan et al., 1998). However, the main difference between other methods and the proposed method is that we tried to bring out details in too dark and bright areas with one-pass processing.

The expected advantages of the proposed scheme are: (1) the algorithm processes fast because it operates directly on the compressed domain; (2) it improves details in dark area and in bright area without any iteration over previous approaches; (3) the algorithm does not affect the compressibility of the original image; and (4) the proposed image enhancement algorithm can be applied to any DCT-based image compression standard, such as JPEG, MPEG, and H.26X without any significant modification.

This paper is organized as follows; in Section 2, we review basic concepts of DCT-based image compression scheme, Retinex algorithm, and a single α-rooting algorithm. Then we present the proposed algorithm in Section 3. In Section 4, we apply the proposed method to different test images to demonstrate and compare its performance with several base-line algorithms. For quantitative comparison of performance, a quality metric based on the perceptual properties of the human visual system and a peak signal-to-noise ratio (PSNR) are used. Section 5 concludes the paper with a brief summary.

2. Background

We review the basic concepts used in this paper. First, we provide a brief introduction to image compression based on DCT. Next, The Retinex theory is described to set the foundation for the dynamic range compression problem. Finally, a single α-rooting is described in relation to the contrast enhancement problem.

2.1. DCT-based compression

We provide a simple synopsis of the JPEG image standard as a DCT-based image compression scheme. It can be easily extended to other DCT-based image/video schemes including MPEG and H.26X.

A base-line DCT-based compression consists of tree basic steps. The first step consists of separating an image into sets or "tiles" of 8×8 pixels. These tiles are then transformed into the spatial frequency domain via the forward DCT. The two-dimensional DCT of an 8×8 block $I_{i,j}$ in an original image is

$$d_{u,v} = \frac{1}{4} C_u C_v \sum_{i=0}^{7} \sum_{j=0}^{7} \cos\frac{(2\mathrm{i}+1)u\pi}{16} \cdot \cos\frac{(2\mathrm{j}+1)v\pi}{16} \cdot I_{i,j}, \tag{1}$$

where

$$C_\eta = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } \eta = 0, \\ 1 & \text{otherwise.} \end{cases} \tag{2}$$

The element in the upper left corner $d_{0,0}$ of a DCT encoded block is called the *DC* coefficient, and the 63 other elements are called *AC* coefficients. The DC coefficient represents the average of a single block in the original image, and the AC coefficients represent variations in gray values in certain directions at certain rates. Fig. 1 shows an 8×8 DCT basis functions. From Eq. (1) and Fig. 1, it can be seen that each coefficient $d_{u,v}$ represents the contribution corresponding to the *uv*th waveform (Bhaskaran and Konstantinides, 1995; Gordon, 1989) and the basis functions corresponding to the coefficients $d_{u,v}$ are arranged in order of increasing spatial frequencies from the upper left to the lower right corner in the horizontal and vertical spatial dimensions, respectively. The second step is required to quantize these blocks with quantization coefficients and to round the block coefficients to the nearest integer values. The greater the amount of compression desired, the coarser the quantization matrix and the greater the number of high frequency