Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 771–777
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Transform coding of monochrome images with a statistical design of experiments approach to separate noise

R. Krishnamoorthi *

Department of Information Technology, Bharathidasan Institute of Technology, Bharathidasan University, Trichy 620 024, Tamilnadu, India

Received 14 May 2004; received in revised form 26 July 2006
Available online 11 January 2007

Communicated by K.R. Namuduri

Abstract

A new transform coding technique based on a set of orthogonal polynomials has been proposed in this paper. In the transformed domain, statistical design of experiments approach is used to separate the spatial variation due to discriminable low level features (signals) from the spatial variation due to an unexplained source called noise. The proposed polynomial transform has a low computational complexity because it is configured as an integer transform. The proposed transform also has very good compression ability or separationability. The degree of this separationability can be controlled by specifying the level of the signal-to-noise ratio.

Keywords: Image coding; Orthogonal transform; Signal processing; Edge; Texture

1. Introduction

In recent years, there has been increased research interests in image coding due to its wide range of applications. There are two types of image coding (i) Lossless and Lossy, while lossless compression could produce the exact original image reproduced after the compression ratio is generally less. Still, in some application is desired and hence techniques to obtain a good quality of the reconstructed picture is studied (Molta et al., 1997). In few applications, where the near original is sufficient, lossy compression is preferred that can yield as high as 97–98% of compression ratio. Model-based compression is studied by many researchers (Sarkar et al., 1997; Krishnamoorthi and Seetharaman, 2005; Bilgin et al., 2000). Transform coding is one of the emerging areas of lossy compression. The main objective of transform coding is to compact the information contained in a sequence into a few subsequences so that the number of bits that are required to encode the sequence can be reduced without causing any distortion to the reconstructed image. In the transform coding of images, the gray levels in a sequence are transformed into another sequence in the spatial-frequency domain so that the major portion of the information is contained in only a few elements of the transformed sequence (Jain, 1981). The compression takes place because these few elements are only required to be retained. Many orthogonal transforms have been proposed for the transform coding of images. Besides transform coding techniques, use of statistical models for monochrome image compression is reported in the literature. Image compression based on Texture Synthesis and Gaussian–Markov Random Field model (Chellappa et al., 1985), Non-Gaussian Autoregressive Models (Kadaba et al., 1998), Principal Component Analysis (Clausen and Wechler, 2000), Arithmetic coding (Sarkar et al., 1997) and Full range Autoregressive model (Krishnamoorthi and Seetharaman, 2005) are available in the literature. Also the use of Wavelet Transforms has gained popularity for grey-scale image compression (Wenberger et al., 1996; Bilgin et al., 2000).

* Tel.: +91 431 2407095/2407222; fax: +91 431 2407333.
E-mail address: rkrish@bdu.ac.in

0167-8655/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.patrec.2006.10.009

In the transform coding scheme, the low computational complexity and high compression ability are the two most important criteria for selecting an orthogonal transform. Among the many different kinds of orthogonal transforms that have been proposed for transform coding, the Discrete Cosine Transform (DCT) is considered to be the best. This has been adopted in the JPEG still image compression standard (Wallace, 1992). There are two major shortcomings in Discrete Cosine Transform coding. Firstly, the computational complexity of the DCT is quite high compared to many nonsinusoidal transform such as C-matrix transform (Kwak et al., 1983), high correlation transform and the low correlation transforms (Cham and Clarke, 1986) and integer cosine transform (Cham, 1989; Liang and Tran, 2001). This is because, these nonsinusoidal transforms consist of only integer components. The other shortcoming of discrete cosine transform coding is that the coding procedure for removing the undesirable coefficients of discrete cosine transformation is merely based on an intuitive notion that the higher order coefficients of DCT have a smaller amplitude. This intuitive transform coding approach is also used with the existing nonsinusoidal transforms resulting in a much lower compression ability than the DCT.

The objective of this work is twofold, namely, to develop a orthogonal transform with a lower computational complexity than DCT and supplement the proposed transformation by a statistical framework such that the transform coding is equivalent to statistical hypothesis testing.

We consider images either untextured, textured or consisting of both untextured as well as textured regions. In untextured images, the main information relating to boundaries of objects within an image is carried in terms of edges whereas in textured images the main information relating to objects is carried in terms of discriminable textures. Hence, the objective of our image compression scheme is to compress image data by discarding non-edge and non-textured information. We assume that these non-edge and non-textured part do not have any effective AC component. Hence, if this AC component is separated and finally discarded there will not be much distortion in the reconstruction. For edges or textured part the AC component will be decomposed into some basis components as per different combinations of frequency and orientation within a source bandwidth. Since only a subset of this basis is sufficient to characterize an edge or texture, information relating to this characterizing subset is to be retained only during compression. This loss of information may easily be tolerated if this is white Gaussian noise with zero mean and constant variance σ_η^2.

2. Polynomial frame work

As the observed gray level $I(x,y)$ at any pixel in an $n \times n$ image region, where x and y are two spatial coordinates, may be considered as a random variate, the occurrences of gray levels in that region can be thought of as yields of an n^2 factorial experiment where the two factors are two spatial coordinates considered at different levels. Hence

$$I(x,y) = g(x,y) + \eta(x,y) \tag{1}$$

where $g(x,y)$ accounts for meaningful variation in $I(x,y)$ due to discriminable edges or texture and $\eta(x,y)$ is the spatial variation due to unexpected sources which may be called noise. At different pixels, we consider $\eta(x,y)$ as independent normally distributed variates with zero mean and constant variance σ_η^2.

The meaningful spatial variation $g(x,y)$ can be parameterized in two different forms for two major classes of image regions, namely, edge response in untextured regions and textured response in case of textured regions (Krishnamorthy and Bhattacharya, 1997).

In order to parameterize the spatial variation $g(x,y)$ either in the form of edge or texture we assume that image regions with edges or texture have gray level variations as a class of orthogonal polynomial functions of the two spatial coordinates. Next, we represent each of the two classes, namely, the edge and texture, in terms of amplitude responses of a suitably chosen closed set of polynomial operators. Let $\Phi = \{O_{ij}(x,y),\ i,j = 0, 1, 2, \ldots, n-1\}$ denote the Polynomial basis, where O_{ij} are polynomial operators. As it has been established in (Krishnamoorthy, 1999) that a Φ_L proper subset of Φ is sufficient to represent uniquely any member of any of the two classes C_L edge or texture. We may characterize in the absence of noise the gray level variation of C_L by

$$g_L(x,y) = \sum\sum_{O_{ij}\in\Phi_L} Z_{ij}\cdot O_{ij}(x,y) \tag{2}$$

In the presence of noise the observed gray level variation of C_L may be expressed as

$$\begin{aligned} I(x,y) &= \sum\sum_{O_{ij}\in\Phi_L} Z_{ij}\cdot O_{ij}(x,y) + \sum\sum_{O_{ij}\notin\Phi_L} Z_{ij}\cdot O_{ij}(x,y) \\ &= \hat{g}(x,y) + \hat{\eta}(x,y) \end{aligned} \tag{3}$$

where Z_{ij} is the amplitude response per unit length of the polynomial operators O_{ij}. The Z_{ij} is obtained by convolving the $n \times n$ image region with O_{ij}. Since Φ is a orthogonal basis, $\hat{g}$ and $\hat{\eta}$ in Eq. (3) are linearly independent. It can also be shown that each Z_{ij}: $O_{ij} \notin \Phi_L$ has an expected value zero and variance σ_η^2. Since for an $n \times n$ image region the observations of gray level variation have a total n^2 degrees of freedom of which 1 degree of freedom is accounted for $\sqrt{n}Z_{00}$, the DC component of each of the remaining $n^2 - 1$ independent Z_{ij}^2 is a $\chi^2\sigma_\eta^2$ variate with 1 degree of freedom. The degrees of freedom meant for that the total number of observations less the number of independent constraints imposed on the observations. As per the Statistical Design of Experiments paradigm we may call z_{ij} and $(z_{ij} - Z_{ij}^2)$, respectively, the estimates of orthogonal effect and variation due to the polynomial source O_{ij}.