European Journal of Radiology 55 (2005) 139–145

EJR
EUROPEAN JOURNAL OF RADIOLOGY

www.elsevier.com/locate/ejrad

HVS-based medical image compression

Xie Kai[a,*], Yang Jie[a], Zhu Yue Min[b], Li Xiao Liang[a]

[a] *Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, 200030 Shanghai, China*
[b] *CREATIS—CNRS Research Unit 5515 and INSERM Unit 630, 69621 Villeurbanne, France*

Received 5 July 2004; received in revised form 8 September 2004; accepted 10 September 2004

Abstract

Introduction: With the promotion and application of digital imaging technology in the medical domain, the amount of medical images has grown rapidly. However, the commonly used compression methods cannot acquire satisfying results.
Methods: In this paper, according to the existed and stated experiments and conclusions, the lifting step approach is used for wavelet decomposition. The physical and anatomic structure of human vision is combined and the contrast sensitivity function (CSF) is introduced as the main research issue in human vision system (HVS), and then the main designing points of HVS model are presented. On the basis of multi-resolution analyses of wavelet transform, the paper applies HVS including the CSF characteristics to the inner correlation-removed transform and quantization in image and proposes a new HVS-based medical image compression model.
Results: The experiments are done on the medical images including computed tomography (CT) and magnetic resonance imaging (MRI). At the same bit rate, the performance of SPIHT, with respect to the PSNR metric, is significantly higher than that of our algorithm. But the visual quality of the SPIHT-compressed image is roughly the same as that of the image compressed with our approach. Our algorithm obtains the same visual quality at lower bit rates and the coding/decoding time is less than that of SPIHT.
Conclusions: The results show that under common objective conditions, our compression algorithm can achieve better subjective visual quality, and performs better than that of SPIHT in the aspects of compression ratios and coding/decoding time.

Keywords: Medical image; Image compression; Wavelet transform; The lifting scheme; Human vision system; Contrast sensitivity function

1. Introduction

Digital technology has given a great advantage to the medical imaging area. Medical images, however, require huge amounts of memory, such as computed tomography (CT), and magnetic resonance imaging (MRI). Due to the limitations of storage and transmission bandwidth of the images, the main problem of the technology lies in how to compress a huge amount of visual data into a low-bit rate stream, because the amount of medical image data would overwhelm the storage device without an efficient compression scheme.

Recently, some researchers proposed human vision system (HVS) in the research of image compression. It can visually remove the information in most degree that human vision cannot preserve. Many solutions have been proposed for embedding a model of the human visual system in compression algorithms. While early perceptually tuned image coders were only concerned with the frequency domain behavior of the HVS [1–3], more recent and efficient coders exploit its spatial domain properties too [4–6], in order to dynamically adjust the parameters of the compression algorithm according to the local properties of the image.

Wavelet-based coders have proven ideally suited for embedding complete HVS models, due to the space–frequency localization properties of wavelet decompositions [7]. The HVS model can be embedded either in the quantization stage [4,6] or in the bit allocation stage [5]. In the former approach a perceptually tuned step is computed for quantizing each DWT coefficient; the quantized coefficient are then entropy encoded and transmitted. In the latter approach, a fixed step is used for quantizing all of the DWT coefficient, but for each

* Corresponding author. Fax: +86 213 2261162.
E-mail address: xie_kai2001@sjtu.edu.cn (X. Kai).

0720-048X/$ – see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.ejrad.2004.09.007

quantized coefficient a suitable number of bits is transmitted, according to its perceptual relevance.

Our algorithm takes a novel approach, which embeds an HVS model in the quantization stage of a wavelet-based coder for visually lossless compression of gray-level images. The paper is organized as follows: in Section 2 some background on the lifting scheme and HVS models is provided. Section 3 details the structure of the proposed algorithm. Finally, experimental results are presented in Section 4.

2. Background

2.1. Lifting scheme

The lifting scheme [8,9] is a new method for constructing biorthogonal wavelets. Its origins lie in a method for improving a given wavelet transform to obtain some specific properties. Subsequently, it was extended to a generic method for creating so-called second-generation wavelets.

The lifting scheme is based on the interpolation to condense information. Some of its advantages as opposed to classical wavelets are the generality of the method, ease of implementation, its speed and its applicability to arbitrary length or geometries. The main difference with classical constructions is that it does not rely on the Fourier transform. Therefore, it can be used to construct wavelets in settings where translation and dilation cannot be used, such as for example, wavelets on bounded domains, on curves and surfaces or in case of irregular sampling.

It consists of three main steps: SPLIT, which subsamples the original data into odd and even sets; PREDICT, which finds the wavelet coefficients as the failure to predict the odd set based upon the even; and UPDATE, which updates the even set by using the wavelet coefficients to compute the scaling function coefficients.

2.2. Physiology of vision

Most visual properties of the HVS are not intuitive. Even when they have been characterized by psychophysical experiments; physiological evidence is the only way to understand the phenomenon completely [10,11].

The physiology of human vision includes the eyes and the retina, where vision is initiated, as well as the visual pathways and the visual cortex, where high-level perception takes place. The eyes represent the first stage of the HVS. They can be understood as a complicated camera continually in motion, allowing accommodation to different light levels and to objects at various distances. The eyes have certain optical defects such as optical blur and chromatic aberration, but normally these do not affect the rest of the processing chain.

Despite our current knowledge of the HVS, its complexity makes it impossible to construct a complete physiological model [12,13]. Some attempts have been made, but they have been restricted to models of the retina and do not account for higher-level perception. Consequently, HVS models used in image processing are usually behavioral and are based on psychophysical studies.

2.3. HVS models for imaging applications

There are also different approaches to HVS-modeling. The unifying rationale is to account for a number of psychophysical effects [14].

2.3.1. Luminance and color

The first stage in the processing chain of HVS models concerns the transformation into an adequate perceptual color space, usually based on opponent colors. After this step, the image is represented by one achromatic and two chromatic channels carrying color difference information. This stage can also take care of the so-called luminance masking or lightness non-linearity [15], the non-linear perception of luminance by the HVS. Such a non-linearity is inherent to more sophisticated color spaces like CIE L*a*b*, but needs to be added to simple linear color spaces. In compression applications, it can be considered by setting the quantization precision of the transform coefficients.

2.3.2. Multi-channel decomposition

It is widely accepted that the HVS bases its perception on multiple channels that are tuned to different ranges of spatial frequencies and orientations. Measurements of the receptive fields of simple cells in the primary visual cortex revealed that these channels exhibit approximately a dyadic structure. This behavior is well matched by a multi-resolution filter bank or a wavelet decomposition. An example for the former is the cortex transform, a flexible multi-resolution pyramid, whose filters can be adjusted within a broad range. Wavelet transforms, on the other hand, offer the advantage that they can be implemented in a computationally efficient manner by a lifting scheme.

It is believed that there are also a number of channels processing different object velocities or temporal frequencies. These include one temporal low-pass and one, possibly two, temporal band-pass mechanisms in the human visual system [16,17], which are generally referred to as sustained and transient channels, respectively.

2.3.3. Contrast and adaptation

The response of the HVS depends much less on the absolute luminance than on the relation of its local variations to the surrounding background, a property known as Weber–Fechner law [15]. Contrast is a measure of this relative variation, which is commonly used in vision models. While it is quite simple to define a contrast measure for elementary patterns, it is very difficult to model human contrast perception in complex images, because it varies with the local image content. Furthermore, the adaptation to a specific luminance level or color can influence the perceived contrast.