Available online at www.sciencedirect.com
SCIENCE DIRECT
Pattern Recognition 39 (2006) 1492–1508
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

A new algorithm for the embedding of a prediction mechanism into the JPEG2000 coding chain

Marco Aguzzi*, Maria Grazia Albanesi, Marco Martelli

Dipartimento di Informatica e Sistemistica, Università degli Studi di Pavia, Via Ferrata 1, 27100 Pavia, Italy

Received 20 July 2005; received in revised form 8 February 2006; accepted 10 February 2006

Abstract

In this paper, we propose the embedding of a prediction mechanism into a part of the coding structure of JPEG2000 image compression standard, in order to reduce the amount of bits sent to the arithmetic coder, without any significant changes into the standard architecture and without loosing performance. The prediction is based upon an innovative processing of the data structures used by the standard JPEG2000 in progressive coding and the addition of a Prediction Matrix, whose computation does not add any overhead at the decoder side. Experiments are performed to test the efficacy of the prediction mechanism, and results are compared to the standard JPEG2000 and other similar approaches. Tests are documented over a set of well-known images from literature, also against different kinds of added noise. Performance, in terms of saved bits are reported, and a new figure of merit is defined to test the efficiency of Prediction. The results prove that the new proposal overcomes the standard and other related approaches for the entire set of referenced images, with significant gain in synthetic images, also in presence of noise.

Keywords: JPEG2000; Prediction; Wavelet transform; Image compression

1. Introduction

With the increasing use of digital media, both in multimedia processing and transmission over networks, the efficient storing and compression of data is a key issue in the current research fields. In particular, the new JPEG2000 [1–5] standard for still image coding is a promising framework, as it overcomes some limitations of the existing standards, such as JPEG [6–8]. In fact, thanks to the features of its encoding scheme EBCOT [9], the use of multiresolution of wavelet decomposition, the context-based arithmetic coding and the rate-distortion optimization, the new standard is able to assure several interesting features: (a) lossy, visually lossless and lossless compression with a better image quality (both from an objective and a subjective point of view), (b) progressive coding and decoding, (c) region of interest definition, and (d) robustness to bit errors. However, despite of its innovative features and high performance, research activities in the field of still image compression are far from being concluded. In particular, new efforts are currently under development [10]: the study of efficient wavelet coefficient modeling [11], the investigation of a useful application-based definition of ROIs [12–14], the analysis of the contexts to enhance the arithmetic coder [15]. In this paper, we investigate the problem of reducing the amount of bits produced by some phases of the stream generation with the following aims: (a) without affecting the global architecture of the standard, (b) without loosing compression performance, and (c) without dramatically increasing the complexity of the decoder.

Our approach consists of embedding a Prediction mechanism in the JPEG2000 coding step called cleanup pass. Prediction is a well-known idea in image coding and transmission; in our case, the prediction is performed in such a way that the prediction error is always kept to zero and the complexity of the realization is kept low, as it requires only the construction of a Prediction Matrix (linear with the number of the coefficients) from the coder point of view,

* Corresponding author. Tel.: +39 0382985523; fax: +39 0382525638.
E-mail address: marco.aguzzi@unipv.it (M. Aguzzi)
URL: http://orfeo.unipv.it/aguzzi.

0031-3203/$30.00
doi:10.1016/j.patcog.2006.02.015

with the introduction of few new contexts for the arithmetic coder. It is important to underline that the Prediction Matrix has not to be transmitted, but it can be computed again at the decoder side, with a fully respect of the progressing coding paradigm and without any overhead to the transmitted stream. In literature, prediction has been implemented in JPEG2000 only in one contribution so far [16], but it refers to Motion JPEG2000 and in this case, Prediction is substantially and conceptually different, as it exploits only correlations among frames to reduce the amount of memory (frame buffers). We can compare our approach to the work proposed in Ref. [17], where the transition between lossless and lossy compression is made smoother using successive refinements. Moreover, if we compare our approach to the new type of coding transforms adapted from brain activity models [18], we underline that we can embed our Prediction mechanism in JPEG2000 without tampering the transform stage.

This paper is organized as follows. In Section 2, an overview of JPEG2000 is given, in Section 3 our prediction algorithm is presented, while Section 4 is focused on the final version; Section 5 presents the experimental results done in order to test the algorithm and Section 6 ends the paper with conclusions and notice to future works.

2. The JPEG2000 coding chain

This section gives a synthetic introduction to JPEG2000 coding chain, while the reader can refer to existing literature for a more detailed description of the coding phases of this emerging standard [19–25].

JPEG2000 coding chain consists of five main functional blocks, namely: (a) Preprocessing, in which unsigned samples are converted to signed ones, (b) Intercomponent transform, in which the color space is converted from RGB to YCrCb (conversion and colorspace information can be found in Refs. [26,27]), (c) Intracomponent transform, in which image samples are processed by a discrete wavelet transform (extended literature exists on wavelet, a selection could be Refs. [28–36], and specific about JPEG2000 is Ref. [37]), (d) Tier-1 coding, in which wavelet coefficients are converted to symbols using a context-driven coder, and (e) Tier-2 coding, in which symbols produced by the previous functional block are converted into the final bitstream.

Tier-1 coding is part of the EBCOT algorithm, and this section will cover it in order to fully understand the algorithm described in the next section. In this step of the compression chain, wavelet coefficients returned by the Intracomponent transform are divided into 64 by 64 squares (usually), called *code-blocks*, and processed by a bit-plane coder, consisting of three main steps, referred to as *coding passes*: (a) *Significance*, (b) *Refinement*, and (c) *Cleanup* (the code by significance scheme is an idea firstly presented in Ref. [38], which became an actual algorithm with SPIHT [39]). In all of these passes, coefficients are processed by means of *stripes*, which are code-block areas four coefficient high and wide as the code-block itself. It is convenient to address a single column of a stripe as *quadruplet* (see Fig. 1(a)). In the three coding passes, the coefficients are scanned from the most significant bit-plane to the least significant bit-plane, in each bit-plane the scan order is from quadruplet to quadruplet (inside each stripe), as specified in Fig. 1(b).

In order to code coefficients, the JPEG2000 standard uses a semi-arithmetic coder (called MQ-coder [40,41]), that is, it follows the behavior of a standard arithmetic coder only relatively to a specified context: in this way the arithmetic coder can be initialized in a number of different ways according to the specified context. Each of the Significance, Refinement, and Cleanup passes has its own contexts: this is useful because each pass codes wavelet coefficients in a different way. In order to implement our algorithm, two new

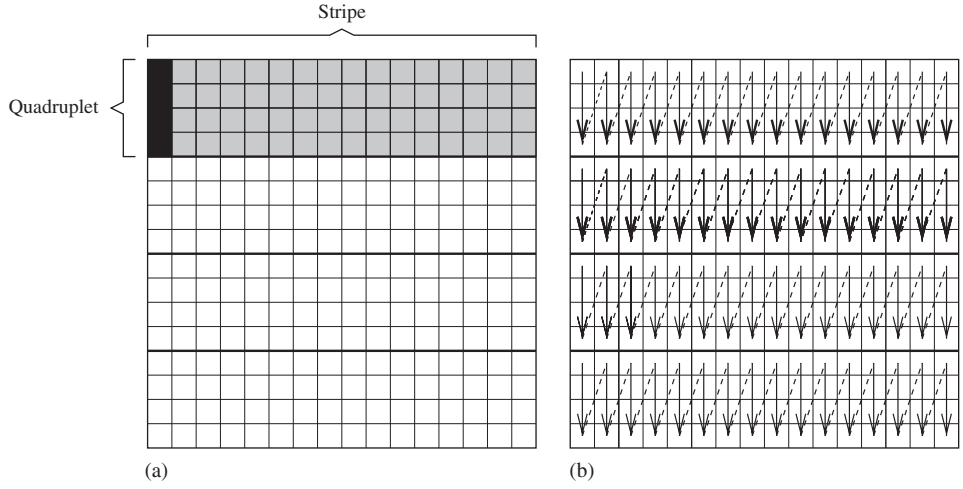

Fig. 1. Single stripe in a code-block and the coefficient scan order: (a) a generic code-block scheme in Tier-1 coding. The stripe is shaded in light gray, the quadruplet in dark gray; (b) scan order of a bit-plane in a code-block.