Digital Signal Processing 30 (2014) 131–140

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Hybrid scheme for efficient shape coding

Zhongjie Zhu [a,*], Gangyi Jiang [b]

[a] Ningbo Key Lab. of DSP, Zhejiang Wanli University, Ningbo 315100, China
[b] Institute of Circuits and Systems, Ningbo University, Ningbo 315211, China

ARTICLE INFO

Article history:
Available online 18 April 2014

Keywords:
Shape coding
Contour-based coding
Near lossless coding
Hybrid scheme
High efficiency

ABSTRACT

A contour-based scheme for near lossless shape coding is proposed aiming to acquire high coding efficiency. For a given shape image, object contours are firstly extracted and then thinned to be perfect single-pixel width. Next they are transformed into chain-based representation and divided into different chain segments based on link directions. Thirdly, two fundamental coding modes are designed and developed to encode different types of chain segments, where the spatial correlations within object contours are analyzed and exploited to improve the coding efficiency as high as possible. Finally, a fast and efficient mode selection method is introduced to select the one that can produce shorter code length out of the two modes for each chain segment. Experiments are conducted and the results show that the proposed scheme is considerably more efficient than the existing techniques.

1. Introduction

Shape is a popular way to define objects and efficient shape coding is a key technique in object-based applications [1–6]. Most of the existing shape coding methods can be mainly classified into two categories: bitmap-based methods and contour-based methods. Bitmap-based methods usually encode a binary shape image directly such as the JBIG [7], JBIG2 [8], and the context-based arithmetic encoding (CAE) technique of MPEG-4 [9], of which MPEG-4 CAE is widely used and many related varieties have been proposed. In [10], an efficient pipelined multi-symbol CAE architecture was proposed for real-time coding. In [11] a generic and scalable shape coding scheme was proposed which can achieve better coding efficiency than non-scalable shape coding techniques. In [12], the rate control issue was studied and an efficient rate-distortion model was proposed.

In recent years, some new bitmap-based techniques have been proposed. A SPIHT-based scheme for coding shape and texture of arbitrarily shaped objects was proposed in [13]. It employs a novel implementation of the shape-adaptive discrete wavelet transform (SA-DWT) using in-place lifting along with parallel coding of texture coefficients and shape mask pixels aiming to create a single embedded code that allows for fine-grained rate-distortion scalability. Objective and subjective simulation results show it has a rate-distortion performance comparable or superior to MPEG-4 CAE. In [14], a block-based scheme that combines a quad-tree structure with context-based arithmetic coding was proposed. Comparing with MPEG-4 CAE, it can considerably improve the coding efficiency. Liu and Ngan explored the H.264/AVC-based shape coding techniques and developed a new scheme that can encode arbitrarily shaped object [15]. In [16], by exploiting the features of locally digital straight line segments (DSLS) contained in binary images, a new DSLS-based coder ($DSLSC_2$) for lossless shape coding was proposed. It is based on a local analysis of the digital straightness of the causal part of object boundary, which is used in context definition for arithmetic encoding. Experimental results show that it outperforms most of the existing techniques such as the JBIG, JBIG2, and MPEG-4 CAE [16].

Generally, bitmap-based methods are fast and simple and can be used for both lossy and lossless shape coding. However, their coding efficiency usually is not high. Unlike bitmap-based techniques, contour-based methods only extract and encode the contours of objects and their coding efficiency are usually high. Typical methods of this kind include the chain coding techniques [17], the curve fitting techniques such as the polygon approximation techniques and the spline approximation techniques [18,19]. Generally, the curve fitting techniques can acquire higher compression efficiency than both the chain coding techniques and the bitmap-based techniques. However, both the chain coding techniques and the bitmap-based techniques can be used for both lossy and lossless applications, while the curve fitting techniques can only be used for lossy purposes.

* Corresponding author.
E-mail address: zhongjiezhu@hotmail.com (Z. Zhu).

http://dx.doi.org/10.1016/j.dsp.2014.04.007

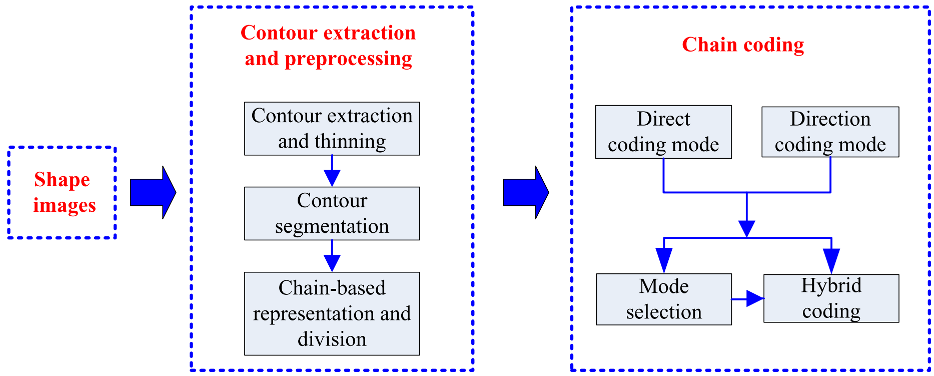

Fig. 1. The diagram of the proposed scheme.

Recently, some new contour-based methods have also been proposed. Wulandhari and Haron proposed a shape coding scheme that can deal with shapes with holes based on vertex chain code [20]. Lai et al. addressed the issue of distortion measurement for B-spline-based shape coding and proposed a new metric that is fairly perceptually consistent [21]. In [22], an attempt to perform shape coding using a multiple grid chain code was proposed, which can support both lossless and quasi-lossless coding. In [23], an image-dependent shape coding algorithm was proposed, which encodes shape as dependent meta data. Both the coding and the decoding processes are designed to be dependent on the underlying image. The correlation between image and shape can be effectively removed and the coding efficiency is improved on average by 300% over most of the existing techniques such as the JBIG, JBIG2, and MPEG-4 CAE. However, this is also a lossy coding scheme and cannot be used for lossless applications.

Since shape information is very important for an object, lossless or near lossless shape coding is desired and expected in many applications. For example, in object-oriented image retrieval applications, since shape is a key feature of objects, the accuracy of shape greatly impact the retrieval efficiency and accuracy. In medical image applications, accurate shape of an organ object can also help to diagnose a disease or undergo an operation. Hence, high efficient lossless or near lossless shape coding is worth being further studied.

In this paper, inspired by the existence of local spatial correlations within object contours as exploited in [16] and the features of chain coding techniques, a high efficient chain coding scheme by exploiting the local spatial correlations within object contours is proposed for near lossless shape coding. For a given binary shape image, we first extract the object contours and thin them to be perfect single-pixel width. Then they are segmented into several contour segments in order that there exists only one path between any two neighboring contour points within a contour segment. For each segment we further transform it into a chain-based representation and divide it into several chain segments according to the link directions to make each chain segment have up to two kinds of links. When encoded, each link can be represented by only one bit. Hence, for a chain segment if we encode its length together with the data of all links, it can be individually represented and encoded. We refer to this kind of coding mode as “direction coding mode”. If a chain segment is long, the coding efficiency under this mode will be high. However, if a chain segment is short, the coding efficiency may not be efficient due to the need and the cost to encode its length. Therefore, in our scheme, another coding mode, direct coding mode, is proposed to deal with this scenario. In the direct coding mode, a chain segment is encoded link by link directly without the need to encode its length. Hence, when a chain segment is short, this coding mode will likely produce a higher coding efficiency than the direction coding mode.

The main contributions of this paper lie in that a new contour-based scheme for near lossless shape coding is proposed where two fundamental chain coding modes are designed and developed to efficiently exploit the spatial correlations within object contours. The major merit of our scheme is that, comparing with most of the existing techniques, the compression efficiency can be greatly improved. The remainder of the paper is organized as follows: Section 2 describes the details of the proposed scheme; partial experimental results and some discussions are given in Section 3; Section 4 concludes the paper.

2. The proposed scheme

The diagram of the proposed scheme is shown in Fig. 1, which mainly consists of two major steps: contour extraction and preprocessing and chain coding. At the first step, the object contours are firstly extracted and thinned to be perfect single-bit width. Then they are segmented into several contour segments based on contour features. Next each contour segment is transformed into chain-based representation based on the Freeman chain code. Following that, to acquire high coding efficiency, the chain of each contour segment is further divided into a number of chain segments according to the directions of chain links. At the second step, two coding modes, direction coding mode and direct coding mode, are firstly developed. Then for each chain segment, the one of the two modes that can produce shorter code length is selected as its final coding mode. The details of the proposed scheme are introduced below.

2.1. Contour extraction and preprocessing

2.1.1. Contour extraction and thinning

There are many techniques that can be used to define and detect object contours. In this paper, we define an object's contour as the set of pixels that belong to the object and have at least one pixel that belongs to the background within their 4-connected neighborhood. As an example shown in Fig. 2, the set of pixels a–g is detected and defined as an object contour, where the white points and the black points denote the foreground object pixels and the background pixels respectively. If $f(x, y)$ is an original binary shape image and the values of the foreground object pixels and the background pixels are 0 and 255 respectively, then the contour extraction can be formulated as

$$C(x, y) = \begin{cases} 0 & \left| f(x, y) \right| + \left| \prod_{p,q} \left| f(p, q) - 255 \right| \right| = 0 \\ 255 & \text{otherwise} \end{cases} \tag{1}$$

where pixels $f(p, q)$ belong to the 4-connected neighborhood of $f(x, y)$. $C(x, y)$ denotes the extracted binary contour map. If $C(x, y) = 0$, it means pixel $C(x, y)$ is a contour point.