EDGE-BASED IMAGE SYNTHESIS MODEL AND ITS APPLICATION TO IMAGE CODING

Makoto Nakashizuka and Hisakazu Kikuchi

† Tokyo University of Agriculture and Technology, 2-24-16 Nakacho, Koganei, Tokyo, 184-0012, Japan
†† Niigata University, 8050 Ikarashi-2, Niigata, 950-21, Japan

ABSTRACT

In this paper, a new image synthesis model based on the wavelet bases is proposed. In the proposed model, images are approximated from the sum of synthesis functions that are shifted to image edges. By applying the proposed model to sketch-based image coding, any iterative image recovery procedure is not required for image decoding. In the design of the synthesis functions, we define the synthesis functions as linear combinations of the wavelet bases. The coefficients for each wavelet base are yield from the iterative procedure. The vector quantization is applied to the vectors of the coefficients to limit the number of the synthesis functions. We apply the proposed synthesis model to the sketch-based image coding. In coding experiment, we show the image coding by eight synthesis functions and comparison with base-line JPEG.

1. INTRODUCTION

Sketch-based image coding is one of the very low-bit rate image compression techniques [1-5]. In the sketch-based image coding [1, 3, 5], images are represented in the form of the contour geometry and intensity differences across the contours. In decoding, images are recovered by the iterative procedure. The iteration minimizes a cost function that is defined by a constraint to the smoothness in intensity changes on planar regions.

The wavelet maxima representation [4] is employed for another approach to the image coding based on image contours. If the wavelet maxima of which basic wavelet function corresponds to the first-order derivative of a smoothing function, the maxima indicate the positions of contours and describe a multiscale behavior of edges on contours. In image recovery from the wavelet maxima, two convex sets are defined [4]. The procedure of image recovery corresponds to the successive projection between two convex sets [6], iterative procedures are required to recover the original images from above both edge-based image coding.

In this paper, we propose the new edge-based image synthesis model. Images are split into two-parts: low-passed image and the high-frequency components around contours of images. The high-frequency components are reconstructed by the sum of synthesis functions that are shifted to the contour positions. The image can be reconstructed by only additions between the low-passed image and the synthesis functions along image contours. It is expected that the computation costs for image decoding will be smaller than the other edge based image coding. We refer the proposed image synthesis model as edge-based image synthesis model. In Chapter 2, we propose the image synthesis model and the design of synthesis function based on the wavelet bases. In Chapter 3, we apply the proposed model to sketch-based image coding.

2. IMAGE SYNTHESIS MODEL AND ITS SYNTHESIS FUNCTION DESIGN

In this paper, we define the edge-based image synthesis model as follow:

$$f(n) = \sum_{m} q(m)g(n-m) + \sum h_{m}(n-m)$$

where $m, n$ are two-dimensional coordinates and consist a pair of two integers. $C$ denotes the set of all contour positions of the image. $N$ denotes the set of all pixel positions of $f(n)$. The image $f(n)$ is approximated by the sum of functions, $\{g(n-m)\}_{m \in \mathbb{N}}$ which are obtained by the linear combination of $\{h_{m}(n-m)\}_{m \in \mathbb{N}}$ which exist only around the contour positions of the image. If $g(n)$ is defined as a smoothing function, the first component of the right hand of the model (1) corresponds to the low-passed image. The high frequency components of edges are recovered by the second component of the model (1). The entire image is approximated by the low-passed image, the contour positions in $C$ and the functions $h_{m}(n)$.

Now, let assume that all $h_{m}(n)$ belong to the set $S$ which consists functions of which number is limited. If the number of the functions in $S$ and the number of the contour positions are small enough, the image can be represent by the small amount of information and the image compression can be achieved. In this paper, the functions which is included in $S$ are referred as synthesis functions. Next, we describe the design method of the synthesis functions for image compression.

Let assume that the all functions $h_{m}(n)$ can be obtained from linear combinations as:

$$h_{m}(n) = \sum_{j=1}^{J} d_{m}(j)\varphi_{j}(n)$$

By this assumption, the model (1) can be represent as the form:

$$f(n) = \sum_{m} q(m)g(n-m) + \sum_{j} d_{m}(j)\varphi_{j}(n-m)$$

To get the coefficients $\{d_{m}(j)\}_{m \in \mathbb{N}}$ which are allocated to each contour position, we represent the each component in Eq. (3) as the vector

$$f = [f(0,0) \cdots f(K-1,0) \cdots f(K-1,K-1)]^T$$

$$f = [f(0,0) \cdots f(K-1,0) \cdots f(K-1,K-1)]^T$$

$$f = [f(0,0) \cdots f(K-1,0) \cdots f(K-1,K-1)]^T$$

$$f = [f(0,0) \cdots f(K-1,0) \cdots f(K-1,K-1)]^T$$
and

\[ d = [d_0(m_0) \cdots d_{j-1}(m_0) \cdots d_0(m_t) \cdots d_{j-1}(m_t)]^T \]

\[ q(0,0) \cdots q(K-1,0) \cdots q(K-1,K-1)^T \]

\( m_t \) is the position which is included in \( C \). \( K \) is the number of pixels of the image along horizontal or vertical direction. We refer the vector \( d \) as the coefficient vector for the image. The approximation \( f' \) which is yield form Eq. (3) can be expressed by

\[ f' = H d \]  

(5)

where \( H \) denotes the transform matrix which consist the column vector corresponds to \( g(n-m) \) or \( \chi(n-m) \). The approximation residue is defined as the difference between the approximation and the original image as:

\[ r = f - H d \]  

(6)

To get the coefficient vector which minimize the squared norm of \( r \), we employ the following iterative operation:

\[ r_i = f - H d_i \]  

(7a)

\[ d_i = d_{i-1} + V r_{i-1} \]  

(7b)

\( H d_i \) and \( r_i \) are the approximation image and the residue after \( i \)-th iteration. \( r \) can be represent as:

\[ r_i = (I - HV) r_{i-1} \]  

(8)

So, If \( V \) is defined as the matrix which satisfies

\[ \| (I - HV) x \| \leq \| x \| \]  

(9)

where \( \| \cdot \| \) denotes the square norm of a vector for any vector \( x \), the squared norm of \( r \) decreases until convergence. Since the rank of \( H \) depends on the number of the contour points and the functions \( \chi(n-m) \), the approximation residue \( r \) does not always converges to zero. Although quasi-optimum coefficient vector \( d \) which decrease the norm of the approximation residue can be obtained by the iteration operation.

In design of the synthesis function, we employ the wavelet bases as \( \chi(n) \). Now, the discrete dyadic wavelet transform is defined as the form of a matrix as:

\[ W = \begin{bmatrix} \psi_{0,0} & \psi_{1,0} & \cdots & \psi_{0,(K-1),(K-1)} & \cdots & \psi_{j-1,0} & \cdots & \psi_{j-1,(K-1),(K-1)} \\ \Phi_{0,0} & \cdots & \Phi_{0,(K-1),(K-1)} & \cdots & \Phi_{j-1,0} & \cdots & \Phi_{j-1,(K-1),(K-1)} \end{bmatrix}^T \]

(10)

By this definition, The wavelet transform of an image \( f \) at \( j \)-th scale, position \((k,l)\) is derived by

\[ w_{j,(k,l)} = [\psi_{j,(k,l)} \cdots \psi_{j,(K-1,K-1)}] f \]  

(11)

\( \Phi_{m,n} \) correspond to the scaling functions of the wavelet transform. The inverse-wavelet transform that corresponds to \( W \) is defined as

\[ \hat{W} = [X_{0,(0,0)} \cdots X_{0,(K-1,K-1)} \cdots X_{j-1,(0,0)} \cdots X_{j-1,(K-1,K-1)}] \]

(12)

If the pair of the course and inverse transform achieve the perfect reconstruction, the pair of two matrices satisfies

\[ I = \hat{W} W \]  

(13)

The transform matrix \( H \) in Eq (5) is defined by the column vectors of the synthesis wavelet bases in Eq. (12) as:

\[ H = [X_{0,0} \cdots X_{0,(K-1)} \cdots X_{0,(K-1),1} \cdots X_{j-1,0} \cdots X_{j-1,(K-1),1}] \]

(14)

This matrix consists only the column vector which corresponds to the wavelet bases which derive the wavelet transform at the contour positions. The matrix \( V \) for the iteration is defined as:

\[ V = [\psi_{0,0} \cdots \psi_{1,0} \cdots \psi_{0,(K-1)} \cdots \psi_{j-1,0} \cdots \psi_{j-1,(K-1)}] \]

(15)

Next, we define the \( \hat{H} \) and \( \hat{V} \) as the matrices which consist the vector of \( W \) and \( \hat{W} \) except from \( H \) and \( V \). Since the assumption of perfect reconstruction, \( HV \) and \( \hat{H} \hat{V} \) satisfies

\[ x - HVx = \hat{H} \hat{V}x \]  

(16)

for any vector \( x \). Since the discrete-dyadic wavelet is a tight frame [7] and the pair of \( \hat{H} \) and \( \hat{V} \) lose the wavelet coefficients during the transform, the relation of squared norm is expressed as:

\[ \| \hat{H} \hat{V}x \| \leq \| x \| \]  

(17)

for any vector \( x \). So, the pair of \( H \) and \( V \) obtained from wavelet transform satisfies the Eq. (9). Since the iteration operation can be achieved by the pair of course and inverse wavelet transforms, the computation costs for an operation of the iteration corresponds to the computation cost of \( I \)-th scale two-dimensional filter banks.

Next, we apply the vector quantization to the coefficient vector to limit the number of the synthesis functions. The partial coefficient vectors

\[ \hat{d}_m = [d_0(m), d_1(m) \cdots d_{j-1}(m)]^T \]

are defined to each contour position \( m \). VQ is now applied for all set of the partial coefficient vectors. Since the each wavelet base is not orthogonal to others, it is difficult to minimize the approximation error for the original image by VQ. So, VQ is done to minimize the error defined as

\[ E = \sum \| d_m - Q(d_m) \| \]

where \( Q(\cdot) \) denotes the vector after quantization. After VQ, the \( h_n(n) \) is approximated to the synthesis functions which defined
3. APPLICATION TO IMAGE CODING

In section 2, we describe the image approximation by the form of (1). The entire image is represent as the linear combination of the smoothed functions and the synthesis functions of which number is limited by the code-book. In this chapter, we apply the edge-based image synthesis model to the sketch-based image coding. We employ the two-directional discrete dyadic wavelet transform for images. In the synthesis function design, we define the wavelet bases as the cubic spline wavelet [4]. Since the zero-crossing point of cubic spline wavelet only exist on the center of entire function, the illegal oscillation will not occur around the contours of the decoded image.

The image contours are detected from the wavelet transform at scale $j=2$ of the original image (Fig. 1). Horizontal and vertical contours are shown in Fig. 2(a) and (b). In contour detection, we set the threshold on the basis of length and modulus maxima of the wavelet transform. The synthesis functions are obtained from the contour positions in Fig. 2(a) and Fig. 2(b). We apply the LGB algorithm to the partial coefficient vectors. The number of representative vector is set as eight for each the horizontal and the vertical edges. The synthesis functions that obtained by VQ are shown in Fig. 3.

In encoding the contour geometry, we employ the chain coding and run-length Huffman coding. The indexes of the synthesis functions are coded by run-length Huffman coding along the contour coding. The low-passed image of the first component of the model (1) is down-sampled by factor 8 and is coded by DPCM. Fig. 4 shows the coding result by the proposed model. Table 1 shows the amounts of coded data for the low-passed image, the contour positions and vector indexes. The bit-rate for the 256x256 image in Fig. 1 is 0.24 bpp including the code-book of VQ. The decoding of the image is realized by only addition between eight synthesis functions and the decoded low-passed image. The JPEG image at same bit-rate is shown in Fig. 5. Comparing with JPEG, image textures and small details have disappeared by the proposed method. However any distortions are not produced. Especially, the contour is well perceived in the decoded image.

<table>
<thead>
<tr>
<th>Table 1 Data amount of code for Fig. 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coarsest approximation images (in bytes)</td>
</tr>
<tr>
<td>Contour positions (in bytes)</td>
</tr>
<tr>
<td>Code-book of VQ (in bytes)</td>
</tr>
<tr>
<td>Index allocated to contour positions (in bytes)</td>
</tr>
<tr>
<td>Total amount of code</td>
</tr>
</tbody>
</table>

The image contours are detected from the wavelet transform at scale $j=2$ of the original image (Fig. 1). Horizontal and vertical contours are shown in Fig. 2(a) and (b). In contour detection, we set the threshold on the basis of length and modulus maxima of the wavelet transform. The synthesis functions are obtained from the contour positions in Fig. 2(a) and Fig. 2(b). We apply the LGB algorithm to the partial coefficient vectors. The number of representative vector is set as eight for each the horizontal and the vertical edges. The synthesis functions that obtained by VQ are shown in Fig. 3.

In encoding the contour geometry, we employ the chain coding and run-length Huffman coding. The indexes of the synthesis functions are coded by run-length Huffman coding along the contour coding. The low-passed image of the first component of the model (1) is down-sampled by factor 8 and is coded by DPCM. Fig. 4 shows the coding result by the proposed model. Table 1 shows the amounts of coded data for the low-passed image, the contour positions and vector indexes. The bit-rate for the 256x256 image in Fig. 1 is 0.24 bpp including the code-book of VQ. The decoding of the image is realized by only addition between eight synthesis functions and the decoded low-passed image. The JPEG image at same bit-rate is shown in Fig. 5. Comparing with JPEG, image textures and small details have disappeared by the proposed method. However any distortions are not produced. Especially, the contour is well perceived in the decoded image.
4. CONCLUSIONS

We propose a new image synthesis model based on the wavelet transform. An image is approximated by the synthesis functions that are shifted to the contour positions. We also propose the synthesis function design by the iterative procedure using the pair of the course and inverse wavelet transform. We apply the proposed model to the sketch-based image coding. By the proposed model, only the additions of the synthesis functions are required for the image synthesis.

The compression results in our study are limited in extremely high compression ratios. Textures and fine structure of images are almost removed in the encoding process. By encoding the removed components, a layered image coding will be obtained. The proposed image coding could play a roll found in the layered image coding by three components: the smoothed image, contours and textures [5].

In this paper, we apply two optimization methods for design of the synthesis function. The first is the iterative procedure that is shown in Eq. (7). The second is LBG algorithm for the partial coefficient vectors. There exists the danger trap to the local minimum. The optimum synthesis function design is also further subject.

REFERENCES