A WAVELET BASED STEREO IMAGE CODING ALGORITHM

Qin Jiang[†], *Joon Jae Lee*[‡] *and Monson H. Hayes III*[†]

[†]Center for Signal and Image Processing, School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, U.S.A.
[‡]Department of Computer Engineering, Dongseo University, Pusan 617-716, South Korea.

ABSTRACT

Stereo image pair coding is an important issue in stereo data compression. A wavelet based stereo image pair coding algorithm is proposed in this paper. The wavelet transform is used to decompose the image into an approximation image and three edge images. In the wavelet domain, a disparity estimation technique is developed to estimate the disparity field using both approximation image and edge images. To improve the accuracy of estimation of wavelet images produced by the disparity compensation technique, a novel wavelet based Subspace Projection Technique(SPT) is developed. In the SPT, the block dependent subspaces are constructed using block varying basis vectors that are derived from the disparity compensated wavelet images. Experimental results show that the proposed algorithm is efficient to achieve stereo image compression.

1. INTRODUCTION

Stereo images are pairs of images acquired from two slightly different perspectives. One can perceive stereo images in three dimension when each image of a stereo pair is viewed by its respective eye of the viewer. Therefore, stereo image pairs provide a two dimensional means to represent three-dimensional scenes. Stereo image pairs have been used in many applications in which the depth perception of the scene is important such as flight simulator for pilot training, medical surgery, virtual reality games, robot vision, military surveillance, and autonomous navigation. With the increase of using stereo images, stereo image compression is important for efficient transmission and storage of stereo images.

There exists an inherent cross-image redundancy between the two images of a stereo pair. By exploiting this redundancy, stereo images can be compressed more efficiently than the two images are independently compressed. A widely accepted coding strategy in stereo image compression is to encode one image independently and then to encode the second image by removing the cross-image redundancy between the two images. Disparity compensation is the most commonly used technique to eliminate the cross-image redundancy within stereo images in stereo image compression [1, 2]. To achieve the compression of stereo images, most existing algorithms encode stereo images in the spatial domain [3, 4, 5]. In this paper, we propose to encode stereo images in the wavelet domain. Specifically, disparity estimation and disparity compensation are performed in the wavelet domain. A novel wavelet based subspace projection technique is developed to improve image estimation over disparity compensation in the wavelet domain. The left image is independently encoded. The right image, on the other hand, is encoded by the disparity compensation and the subspace projection technique in the wavelet domain. The diagram of the proposed algorithm is presented in figure 1. In the following section, the

This work was supported by Joint Services Electronic Program under contract DAAH-04-96-1-0161.

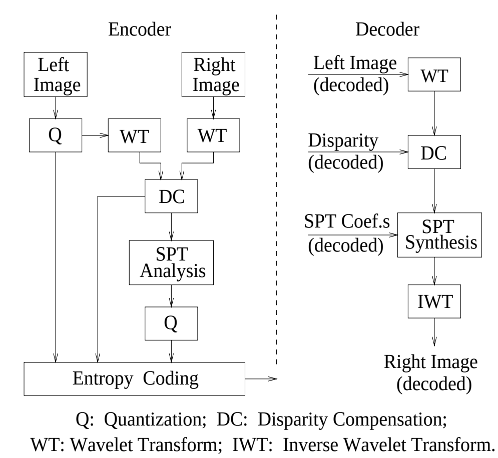

Figure 1: The diagram of wavelet based stereo image coding algorithm

wavelet decomposition is briefly reviewed and the disparity estimation and compensation in the wavelet domain are described. The wavelet based subspace projection technique is presented in section 3. Section 4 presents experimental results on the proposed algorithm. Our conclusion is given in section 5.

2. DISPARITY ESTIMATION AND COMPENSATION IN THE WAVELET DOMAIN

2.1. Wavelet Decomposition for Image Coding

The wavelet decomposition is a multiresolution representation of a signal using a set of basis functions generated by the dilation and translation of a unique wavelet function. Let $\phi(t)$ be a low pass scaling function and $\psi(t)$ be an associated band pass wavelet function. Using separable products of the scaling function $\phi(t)$ and wavelet function $\psi(t)$, a two dimensional wavelet decomposition can be constructed. With the two dimensional wavelet decomposition, an image function $f(x,y)$ can be decomposed into a set of independent, spatially oriented frequency channels [6]. For one level decomposition, the discrete two-dimensional wavelet transform of the image function $f(x,y)$ can be written as

$$Af = ((f(x,y)*\phi(-x)\phi(-y))(2n,2m))_{(n,m)\in Z^2} \quad (1)$$

$$D^1 f = ((f(x,y)*\phi(-x)\psi(-y))(2n,2m))_{(n,m)\in Z^2} \quad (2)$$

$$D^2 f = ((f(x,y)*\psi(-x)\phi(-y))(2n,2m))_{(n,m)\in Z^2} \quad (3)$$

$$D^3 f = ((f(x,y)*\psi(-x)\psi(-y))(2n,2m))_{(n,m)\in Z^2} \quad (4)$$

and an implementation of this decomposition is shown in figure 2 where $h = \phi$ and $g = \psi$.

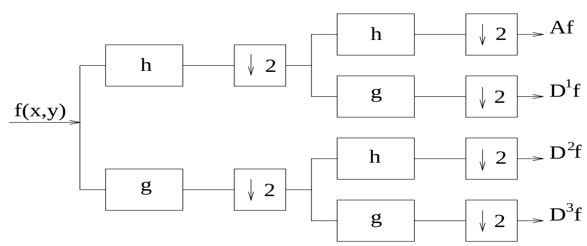

Figure 2: Wavelet decomposition of an image $f(x,y)$

The image function $f(x,y)$ is decomposed into four components. The component Af(LL channel) is an approximation image of $f(x,y)$ and the components D^1f(LH channel), D^2f(HL channel) and D^3f(HH channel) are three detail images of $f(x,y)$. The component Af corresponds to the lowest frequencies, D^1f gives the vertical high frequencies(horizontal edges), D^2f the horizontal high frequencies(vertical edges) and D^3f the high frequencies in the both directions(diagonal edges). Each of the four components can be further decomposed into its four subcomponents. Most energy of the image $f(x,y)$ is concentrated in the wavelet image Af that is only one fourth of the image $f(x,y)$ in size. This is a good property for data compression since the purpose of image data compression is to compactly represent image data.

2.2. Disparity Estimation and Compensation Using Wavelet Representation

Since the wavelet decomposition provides both image intensity patterns in the approximation image and image edge patterns in the detail images within a multiresolution representation, it is useful for image analysis and image estimation. In the research of computer vision, some researches have used the wavelet representation of stereo images to solve *the correspondence problem* (stereo matching) [7, 8]. For stereo image compression, the disparity estimation is important for using disparity compensation to eliminate the cross-image redundancy between the two images of a stereo pair. The following are some advantages to estimate the disparity field using the wavelet representation of the stereo images.

- Edge information such as the length and the orientation of edges that are available in the wavelet domain may be used to improve the estimation of the disparity field.
- The size and the dynamic range of the disparity field are reduced in the wavelet domain because of the down sampling in the wavelet decomposition. This may help to save some bits to transmit the disparity field.
- Using a coarse to fine estimation strategy within multiresolution wavelet representation, computationally, the disparity field can be more efficiently estimated in the wavelet domain than in the spatial domain [8, 9].

In the wavelet domain, since the four wavelet image components represent the same scene, they have the same disparity field. We jointly use the four wavelet image components to compute the disparity field. The four wavelet image components, however, contain different energies; they are not equally important to reconstruct original image from them. To exploit this property in the disparity estimation, the dynamic range of each component is used to weight its distortion in the disparity estimation in this paper. For the sake of simplicity in computation, the proposed disparity estimation technique uses a block based estimation technique and is described as follows.

Let I_r and I_l denote the right image and the left image and $\{AI_r, D^iI_r,\ i=1,2,3\}$ and $\{AI_l, D^iI_l,\ i=1,2,3\}$ denote the wavelet images(decomposition) of I_r and I_l. At first, the wavelet images for the right image are divided into $m \times n$ blocks. Assuming $\mathbf{b}_r^a(\mathbf{x})$ be a block with center at $\mathbf{x}$ in AI_r and $\mathbf{b}_r^i(\mathbf{x})$ be a block with center at $\mathbf{x}$ in D^iI_r. Then, for each block $\mathbf{b}_r^a(\mathbf{x})$, the disparity vector $\mathbf{d}$ is determined by minimizing the functional

$$\begin{aligned} J(\mathbf{d}) &= w^a\|\mathbf{b}_r^a(\mathbf{x}) - \mathbf{b}_l^a(\mathbf{x}+\mathbf{d})\| \\ &+ \sum_{i=1}^{3} w^i \|\mathbf{b}_r^i(\mathbf{x}) - \mathbf{b}_l^i(\mathbf{x}+\mathbf{d})\|, \end{aligned} \quad (5)$$

where $\mathbf{d}$ is searched within a searching window and w^a and w^i are computed by normalized dynamic ranges of corresponding channels. Specifically,

$$\begin{aligned} w^a &= \frac{(max(AI_r) - min(AI_r))}{Q} \\ w^i &= \frac{(max(D^iI_r) - min(D^iI_r))}{Q}, \\ Q &= (max(AI_r) - min(AI_r)) \\ &+ \sum_{i=1}^{3}(max(D^iI_r) - min(D^iI_r)). \end{aligned}$$

Disparity compensation, in general, is to use one image and the disparity field to estimate the second image in stereo image compression. In our approach, the right wavelet image $\{AI_r, D^iI_r,\ i=1,2,3\}$ is estimated using the disparity field and the left wavelet image $\{AI_l, D^iI_l,\ i=1,2,3\}$. Let $\mathbf{b}_{dc}^a$ denote a disparity compensated block, the block $\mathbf{b}_{dc}^a$ is computed by translating the block $\mathbf{b}_l^a$ with the disparity vector $\mathbf{d}$; that is

$$\mathbf{b}_{dc}^a(\mathbf{x}) = \mathbf{b}_l^a(\mathbf{x}+\mathbf{d}) \quad (6)$$

Then, the block $\mathbf{b}_{dc}^a(\mathbf{x})$ is used as an estimate of block $\mathbf{b}_r^a(\mathbf{x})$. Because the four channels share the same disparity field, the blocks in other three channels are estimated in the same way.

3. THE SUBSPACE PROJECTION TECHNIQUE IN THE WAVELET DOMAIN

The principle of the subspace projection is to project a vector onto a subspace. Then, the subspace representation of the vector is used as an estimate of the vector. Generally, the dimension of the subspace is much smaller than the dimension of the vector. This greatly reduces the amount of data to be compressed. The key issue here is how to construct a subspace that can produce a better representation for the vector. For a block based image estimation, different block in an image may have different properties. Apparently, in order to obtain a better representation for the image, the subspace should be able to adapt to the properties of each block in the image. This means that we need to construct a block dependent subspace for each block to achieve a better subspace representation. Since a subspace can be completely determined by its basis vectors, constructing block varying basis vectors is important to the subspace projection technique.

In the wavelet based SPT, we use the disparity compensated wavelet images $\{AI_{dc}, D^iI_{dc},\ i=1,2,3\}$ to construct subspaces for the wavelet images $\{AI_r, D^iI_r,\ i=1,2,3\}$. Let $\{\mathbf{b}_{dc}^0=\mathbf{b}_{dc}^a, \mathbf{b}_{dc}^i,\ i=1,2,3\}$ denote the blocks within $\{AI_r, D^iI_r,\ i=1,2,3\}$. For each block $\mathbf{b}_r^a$, we construct a subspace S by

$$S \;=\; span\{\mathbf{b}_{dc}^i,\ i=0,1,2,3\}, \qquad S \subset \mathbf{R}^{m\times n}. \tag{7}$$

Since $\mathbf{b}_{dc}^i,\ i=0,1,2,3$ come from four orthogonal channels, they are, in general, linearly independent. The vector $\mathbf{b}_{dc}^0$ provides an approximated low frequency image, the vector $\mathbf{b}_{dc}^1$ an approximated horizontal edge image, the vector $\mathbf{b}_{dc}^2$ an approximated vertical edge image and the vector $\mathbf{b}_{dc}^3$ an approximated diagonal edge image. They are block dependent; hence, they are able to adapt to the local properties of the image. The subspace S derived from them may be expected to be able to produce a better representation for the wavelet images $\{AI_r, D^iI_r,\ i=1,2,3\}$. For example, the estimation error between $\mathbf{b}_{dc}^a$ and $\mathbf{b}_r^a$ becomes larger at image edges, which is related to the disparity discontinuity in the image, and this error may be well represented by the three basis vectors containing edge information, $\mathbf{b}_{dc}^i,\ i=1,2,3$. To determine the subspace representation of the block $\mathbf{b}_r^a$ using the subspace S, we can project it onto the subspace S. Assuming P_s is an orthogonal projection operator, $P_s:\mathbf{R}^{m\times n}\rightarrow S$. Then, the projection of $\mathbf{b}_r^a$, denoted by $\mathbf{b}^*$, can be written as

$$\mathbf{b}^* \;=\; P_s\mathbf{b}_r^a \;=\; \sum_{i=0}^{3}\alpha_i\mathbf{b}_{dc}^i, \tag{8}$$

for some $\alpha_i \in \mathbf{R}$.
The projection coefficients, $\alpha_i,\ i=0,1,2,3$ can be computed by the orthogonal projection principle. According to the orthogonal projection principle, the error vector between the vector $\mathbf{b}_r^a$ and its projection vector $\mathbf{b}^*$ is orthogonal to the subspace S. Therefore,

$$(\mathbf{b}_r^a - \mathbf{b}^*) \perp \mathbf{b}_{dc}^i, \quad i=0,1,2,3 \tag{9}$$

and

$$<\mathbf{b}_r^a - \mathbf{b}^*, \mathbf{b}_{dc}^i> \;=\; 0, \quad i=0,1,2,3. \tag{10}$$

Substituting $\mathbf{b}^*$ by the equation (8) gives

$$<\mathbf{b}_r^a, \mathbf{b}_{dc}^i> = \sum_{j=0}^{3}\alpha_j <\mathbf{b}_{dc}^i, \mathbf{b}_{dc}^j>, \quad i=0,1,2,3. \tag{11}$$

If we assume

$$r_{ij} \;=\; <\mathbf{b}_{dc}^i, \mathbf{b}_{dc}^j>, \quad \mathbf{r} \;=\; [r_{ij}],$$

$$c_i \;=\; <\mathbf{b}_r^a, \mathbf{b}_{dc}^i>, \quad \mathbf{c}=[c_0,c_1,c_2,c_3]^\top,$$

the equation (10) can be represented in terms of matrix as follows

$$\mathbf{c} \;=\; \mathbf{r}[\alpha_0,\alpha_1,\alpha_2,\alpha_3]^\top \tag{12}$$

then,

$$[\alpha_0,\alpha_1,\alpha_2,\alpha_3]^\top \;=\; \mathbf{r^{-1}c}. \tag{13}$$

The projection coefficients, $\alpha_i,\ i=0,1,2,3$ have to be encoded then be transmitted over the communication channel. At the decoder, the SPT synthesis procedure reconstructs the block $\mathbf{b}_r^a$ using the quantized projection coefficients, $\alpha_i^q,\ i=0,1,2,3$ and the disparity compensation $\mathbf{b}_{dc}^i$, *i.e.*

$$\hat{\mathbf{b}}_r^a \;=\; \sum_{i=0}^{3}\alpha_i^q\mathbf{b}_{dc}^i, \tag{14}$$

One of advantages to use the wavelet based SPT is that we can only apply the SPT to the channel AI_r and other three channels are estimated simply by the disparity compensation to obtain a better reconstructed image of I_r because most of energy of I_r is concentrated in the channel AI_r. This can save many bits to encode the SPT coefficients comparing with applying the SPT in the spatial domain in which the SPT is generally applied to the whole image.

4. EXPERIMENTAL RESULTS

Three stereo pairs, "Lab", "BookSale" and "Crowd", have been used to test the proposed algorithm. The image sizes are 512×480 for "Lab", 640×240 for "BookSale" and "Crowd". The "BookSale" and "Crowd" are obtained from two AVDS stereo sequences from Carnegie Mellon University. The left images are independently encoded by a subband coding technique with a hight quality and right images are encoded based on the coded left images. The block size used in the block based disparity estimation and the wavelet based SPT is 8×8. The mean absolute difference is used as the distortion measure in the disparity estimation. The disparity fields are encoded by a lossless Huffman coding technique. The the SPT coefficients are encoded by a subband coder to exploit the correlation between the coefficients of neighbor blocks. In our experiments, the shorter the length of the wavelet filter, the higher the PSNR on the reconstructed right images. Hence, the Daubechies wavelet filter with length 2 is used in this paper. Two wavelet based stereo image coding schemes are tested in this paper. One is called WTSIC-I in which the SPT is only applied to the channel Af and other three channels are encoded by the disparity compensation. Another is called WTSIC-II in which the SPT is applied to two channels, Af and $D^1 f$, and the disparity compensation is used to encode the other two channels. In terms of the rate-distortion performance, we compare these two coding schemes with a standard coding technique, the disparity compensation in the spatial domain(DC-SP) with 8×8 block size on all images. The results are reported in table 1. The original right image of "Lab" and a coded right image are presented in figure 3 and figure 4. The experimental results on our test images show that the proposed algorithm performs better than the standard disparity compensation technique in terms of the rate-distortion performance.

Figure 3: Original right image

Figure 4: Coded right image at 0.074bpp with a PSNR of 29.63dB.

5. CONCLUSION

We proposed a wavelet based stereo image coding algorithm in this paper. The wavelet representation of stereo images are used to estimate the disparity field and the disparity compensation is performed in the wavelet domain. A wavelet based subspace projection technique is developed to improve the image estimation over the disparity compensation technique. In the subspace projection technique, the block dependent subspaces are constructed using block varying basis vectors that are derived from the disparity compensated wavelet representation. The experimental results show that the wavelet based stereo image coding algorithm is efficient to achieve the compression of stereo images.

Image Name	**DC-SP**		**WTSIC-I**		**WTSIC-II**	
	PSNR	bpp	PSNR	bpp	PSNR	bpp
Lab	28.84	0.081	29.63	0.073	29.78	0.080
Book	22.57	0.059	23.47	0.055	25.07	0.062
Crowd	21.62	0.051	21.97	0.048	22.77	0.053

Table 1: Comparison of the rate-distortion performance of the proposed algorithm with the disparity compensation in the spatial domain

6. REFERENCES

[1] H. Yamaguchi, Y. Tatehira, K. Akiyama, and Y. Kobayashi, "Stereoscopic image disparity for predictive coding," *IEEE ICASSP*, pp. 1976–1979, 1989.

[2] A. Kopernik and D. Pele, "Improved disparity estimation for the coding of stereoscopic television," *SPIE VCIP*, vol. 1818, pp. 1155–1166, 1992.

[3] H. Aydınoğlu and M. Hayes, "Stereo image coding," *Int. Symposium on Circuits and Systems*, vol. I, pp. 247–250, April 1995.

[4] M. Perkins, "Data compression of stereopairs," *IEEE Trans. on Communications*, vol. 40, No. 4, pp. 684–696, April 1992.

[5] T. Ozkan and E. Salari, "Coding of stereoscopic images," *Image and Video Processing, SPIE*, vol. 1930, pp. 228–235, 1993.

[6] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-11, pp. 674–693, July 1989.

[7] H.-P. Pen, "General stereo image matching using symmetric complex wavelets," *Wavelet Applications in Signal and Image Processin IV, SPIE*, August 1996.

[8] Y. S. Kim, J. J. Lee, and Y. H. Ha, "Stereo matching algorithm based modefied wavelet decomposition process," *Pattern Recognition*, vol. 30, no. 6, pp. 929–952, 1997.

[9] S. Sethuraman, A. Jordan, and M. Siegel, "Multiresolution based hierarchical disparity estimation for stereo image compression," *Symposium on Application of Subbands and Wavelets*, 1994.