Stereo Image Compression by Quadrant Vector Quantization

Chompoonuch Tengcharoen[a] *, Ruttikorn Varakulsiripunth*[b]
Tanasak Phanprasit[c] *and Manas Sanworasil*[d]

[a,b,d]Faculty of Engineering,
and Research Center for Communications and Information Technology (ReCCIT),
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand.
[c]Department of Electronics Engineering, Bangkok Univesity,
Rangsit, Pathumtani, Bangkok 12120 Thailand
E-mail: chompoonucht@yahoo.com[a], kvruttik@kmitl.ac.th[b]
Tanasak.p@bu.ac.th[c], ksamanas@kmitl.ac.th[d]

Abstract

A conventional stereoscopic system with a single left and right image pair that should be transmitted simultaneously is twice of the raw data in a monoscopic image system, so it needs an effective coding algorithm. Disparity compensation is a good methodology for stereo image compression that uses one view as a reference to predict another, and the difference data is coded. Due to the application of reference image, it also needs the effective coding. Therefore, we have proposed a new efficient coding technique, called Quadrant Vector Quantization (QVQ) based on wavelet transform.

Keywords: stereoscopic image, vector quantization, wavelet transform, disparity compensation

1 Introduction

A number of applications can benefit from stereo or multi-view imagery, including tele-presence, remote operation, and entertainment. In general, the stereo images are generated by using two cameras to record two slightly different perspective view point of the same scene, as left and right image. Each image must be viewed by respective eyes, i.e., see the left image by the left eye and see the right image by the right eye. By using special glasses to see them on the monitor screen, the eye-brain will fabricate the stereoscopic depth of the image.

A conventional stereo system with a single left-right pair needs double raw data when comparing with a monoscopic imaging system [1]. And since a stereoscopic image pair essentially depicts the same scene from two different points of view, therefore the independent coding of both images of a stereoscopic pair is redundant [2]. However, it can be reduced the size by using disparity compensation. In this paper, we use the information of left image as the reference to predict the residual (right) image.

Likewise, the reference image also needs an effective coding technique. Hence, we propose a novel algorithm, i.e., the quadrant vector quantization, which is simpler and faster than the general vector quantization (VQ). This is because it uses an intelligent algorithm to create the training vector and searching codebooks for reconstructing the image.

The compression scheme of our proposal consists of encoding and decoding parts. And the components of both parts are shown in Figure 1. The encoder of left and right stereo image is described separately in Section 2.1 and 2.2 respectively. The decoder is described in Section 3. Finally, the implementation and simulation results are analyzed in Section 4.

2 Encoder

Our encoder consists of two processes. Firstly, the left image is compressed by using QVQ techniques. Secondly, the reconstructed left image is used as reference to do disparity compensation with the right image.

2.1 Quadrant Vector Quantization

Figure 2 shows the processes of the quadrant vector quantization (QVQ) of the left image. Firstly, the left image is decomposed by wavelet transform. Secondly, the transformed result is passed to the training vector section in order to generate the suitable codebook. The details of these two sub-processes are described in section 2.1.1 and 2.1.2, respectively.

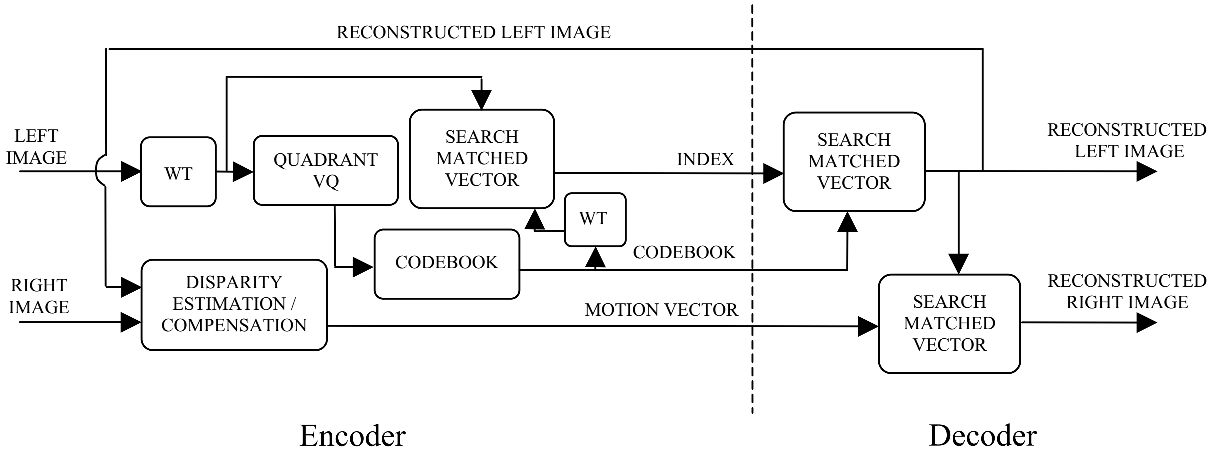

Figure 1: The stereo image compression scheme of our proposal. The left and right parts of dash line are encoder and decoder, respectively.

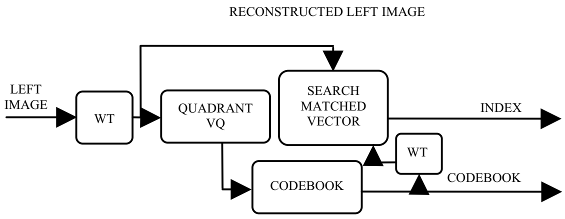

Figure 2: Process of quadrant vector quantization (QVQ) of left image in encoder.

2.1.1 Wavelet Decomposition

As shown in Figure 2, initially, the inputted left image will be divided into small blocks of the same size as 4x4 pixels. The number of divided blocks is in according to the original size of the left image. For example, if the original size of the left image is 64x64 pixels, then we will have 256 divided blocks, and each block's size is 4x4 pixels. Then, each block of the original left image is successively decomposed two times by wavelet transform (WT) as shown in Figure 3 and 4. The wavelet decomposition is a multi-resolution representation of a signal by using a set of basis functions thatgenerated 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. Then, the two-dimensional wavelet decomposition can be constructed by using the separable products of $\phi(t)$ and $\psi(t)$. The discrete two-dimensional wavelet transform of the image function $f(x,y)$ in one level decomposition can be written as follow [5];

$$A_1 f = ((f(x,y) * \phi(-x)\phi(-y))(2n,2m))_{(n,m)\in z^2} \quad (1)$$

$$D_1^1 f = ((f(x,y) * \phi(-x)\psi(-y))(2n,2m))_{(n,m)\in z^2} \quad (2)$$

$$D_1^2 f = ((f(x,y) * \psi(-x)\phi(-y))(2n,2m))_{(n,m)\in z^2} \quad (3)$$

$$D_1^3 f = ((f(x,y) * \psi(-x)\psi(-y))(2n,2m))_{(n,m)\in z^2} \quad (4)$$

The implementation of this decomposition is shown in Figure 3, where $h = \phi$ and $g = \psi$.

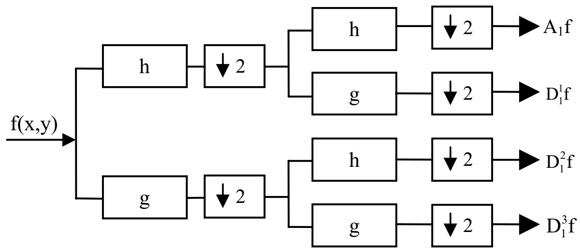

Figure 3: Wavelet decomposition of left image function $f(x,y)$.

As shown in Figure 3 and equations (1), (2), (3) and (4), the image function $f(x,y)$ is decomposed into four components. The component A_1f (LL channel) is an approximation part of image and the other components D_1^1f (LH channel), D_1^2f (HL channel) and D_1^3f (HH channel) are the details of the image. Therefore, most energy of the image $f(x,y)$ is condensed in the wavelet transform image, i.e., A_1f component. In the first level of wavelet decomposition, the individual image block of size 4x4 pixels is decomposed into 4 small blocks of size 2x2 pixels. Therefore, A_1f will be a block of size 2x2 pixels, at the first time. After decomposition in the second level, the smallest wavelet coefficients inside four components are obtained as shown in Figure 4. Each coefficient pixel has its sign, positive or negative, based on wavelet transform. We will use these characteristics for training the vector and then generating compact codebook.

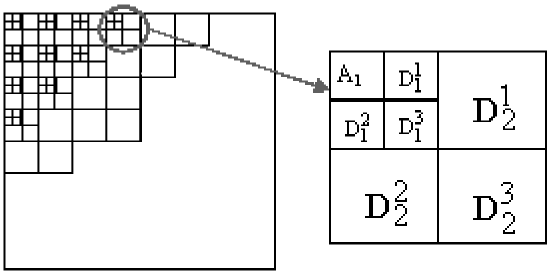

Figure 4: Two times of successive decomposition wavelet transform.

2.1.2 The QVQ training vector

Hence, the approximation A_1f keeps dc energy of its blocks and encoded separately from the details of image. Therefore, only the three detail components of each block are used to create the codebook. The stipulation to create a quadrant codebook is obtained by considering symbolic wavelet coefficients of D_1^1f, D_1^2f and D_1^3f components. Since each coefficient of the component can be expressed in two signs (+/-), then these data types are separated into 2^3 patterns of quadrant codebook as shown in Table 1.

Table 1: 2^3 patterns of the quadrant codebook.

	Sign of coefficient in component		
	D_1^1f	D_1^2f	D_1^3f
Case 1	+	+	+
Case 2	+	+	-
Case 3	+	-	+
Case 4	+	-	-
Case 5	-	+	+
Case 6	-	+	-
Case 7	-	-	+
Case 8	-	-	-

Next, the data in each 4x4 pixels' block is expanded into 1x16 pixels vector size and allotted to be in each appropriate sign quadrant codebook. Here, the training vector is needed to generate most suitable codebook in error admitting level. Our proposed technique can do it very simply by reducing index into only 8 codebook groups in according to its definition. Therefore, the vector is directly arranged in its sign codebook at first, and then the appropriate number of codewords is generated. Finally, each codeword is re-arranged from 1x16 pixels vector to matrix of 4x4 pixels' block. And each block is decomposed successively twice with wavelet transform in order to get the index by searching a matched vector from wavelet coefficient of the original left image. The index of codeword obtained by this process will offer the lowest distortion.

It is clear that the quadrant codebook is generated based on wavelet decomposition and training vector operation. Once the closest codeword is found, the index of that codeword is sent through a channel (the channel can be computer storage, communications channel, and so on) followed by all codebooks. The importance of this method is the creation of the smallest effective codebook that can be used as the reference to reconstruct the satisfied quality image.

2.2 Disparity Compensation

After QVQ process, the reconstructed left image is used as a reference to predict right image. And by using disparity compensation process, the error of reconstructing the right image can be reduced.

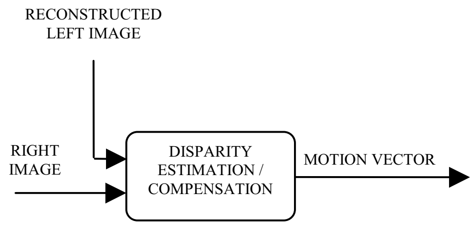

Figure 5: Disparity compensation process.

Figure 5 shows the disparity compensation for encoding right image. Initially, the classical block matching [3] algorithm is employed in order to estimate correspondence between blocks of the two images. The right image is partitioned into 8x8 pixels-block and each block is scanned to match with a corresponding block in the reference feedback reconstructed left image. Here, let **B** represents a partitioned 8x8 pixels-block of right image. Then, the displaced frame difference (DFD) [4] for a block **B** in the right image is defined as follow;

$$\mathrm{DFD}(\mathrm{B}) = \sum_{(x,y)\in \mathrm{B}} \left| I_r(x,y) - I_l(x+v_x, y+v_y) \right| \quad (5)$$

where $I_r(x,y)$ and $I_l(x,y)$ denote the image intensities for the right and left image respectively. Now $\mathrm{v} = \left[v_x, v_y \right]$ [4] is the corresponding disparity vector and the indicated disparity vector for a block is obtained as follow;

$$\mathrm{v} = \arg \min_{v_x, v_y, |v_x| \le \Delta_x, |v_y| \le \Delta_y} \mathrm{DFD}(\mathrm{B}) \quad (6)$$

where Δ_x and Δ_y indicate the limits of the search window.

3 Decoder

Decoder is a part of reconstructing the left and right image as shown in Figure 1. The left image is reconstructed from QVQ process and the right image is reconstructed from disparity compensation process.

For reconstructing the left image, when the decoder receives the codebook and the index of the codeword from encoder part, it replaces the index with the associated codeword. Similar with VQ, the QVQ is a lossy data compression method based on the principle of block coding. According to the explanation in section 2.2, the reconstructed left image is used as the reference to reconstruct the right image by searching match vector with motion vector and compensation. The implementation of the proposed compression technique and its simulation results are described in Section 4.

4 Implementation and Results

We used MATLAB version 6.1 with Pentium4 - 1.70 GHz computer for the simulation. The input sources are four stereo image pairs in bitmap format and their sizes are 64x64 and 128x128 pixels, gray scale images. The synthetic "Room" image represents virtual reality, while the two scenes provide different distance scenarios that exhibit different disparity properties. The left and right image of the "Room" pair differ mainly in the left edge of the left image where a piece of the wall is visible that can not be found in the right image. The "Fruits" are real image that have slightly different rectangles. The "Aqua" pair is occurred the different around the rock in the middle and the "Outdoor" pair has smallest different because of the larger distance between the objects and the cameras.

First, we compress the left image by QVQ with haar wavelet and sent its reconstructed image to compress right image with vector estimation and compensation. The feedback reconstructed left image will reduced error for the reconstructing right image because it is also using in the decoding part. The example of reconstructed right image from disparity compensation process is shown in Figure 6. And the results of reconstruct left image by QVQ coding technique using haar wavelet are represented in Figure 7-9.

The mean peak-signal-to-noise ratio [4, 5] is used as a measure of the stereo pair's reconstruction quality (in decibels) as follow;

$$PSNR = 10\log_{10}\frac{255^2}{\left(D_l + D_r\right)/2} \qquad (7)$$

where D_l and D_r are the difference between the original and reconstructed of left and right image, respectively.

The results of stereo image compression obtained by proposed quadrant vector quantization are represented as bit per pixel (bpp), compression ratio (CR) and Peak Signal to Noise Ratio (PSNR) in Table 2 and 3.

Table 2: Reconstructed stereo image quality in size 64x64 pixels in 8 bits (decibels for PSNR).

Stereo pair	bits/pixel	CR	PSNR
Aqua	0.49929	16.023	33.321
Fruits	0.49858	16.046	34.567
Outdoor	0.49822	16.057	32.819
Room	0.49190	16.264	33.281

Table 3: Reconstructed stereo image quality in size 128x128 pixels in 8 bits (decibels for PSNR).

Stereo pair	bits/pixel	CR	PSNR
Aqua	0.47168	16.961	25.340
Fruits	0.46309	17.275	25.892
Outdoor	0.46723	17.085	24.026
Room	0.48257	16.578	25.260

(a) (b) (c)

Figure 6: (a) and (b) are the original "Room" in left and right stereo images, respectively. (c) is a reconstructed right image from disparity compensation method.

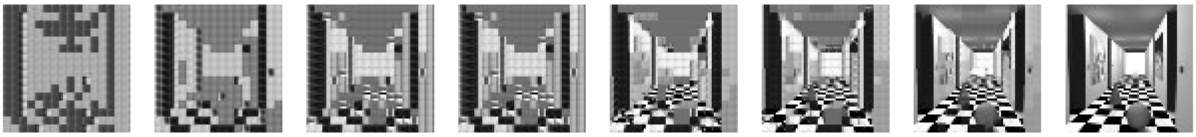

Figure 7: The reconstructed left image "Room" size 64x64 pixels in 1- 8 bits respectively based on QVQ technique using haar wavelet.

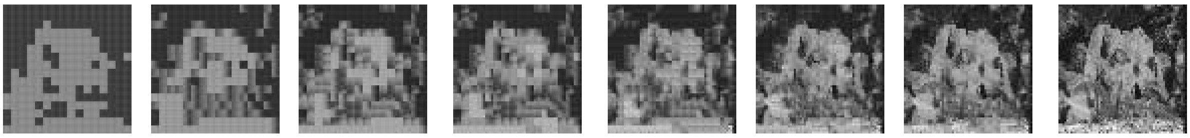

Figure 8: The reconstructed left image "Aqua" size 64x64 pixels in 1- 8 bits respectively based on QVQ technique using haar wavelet

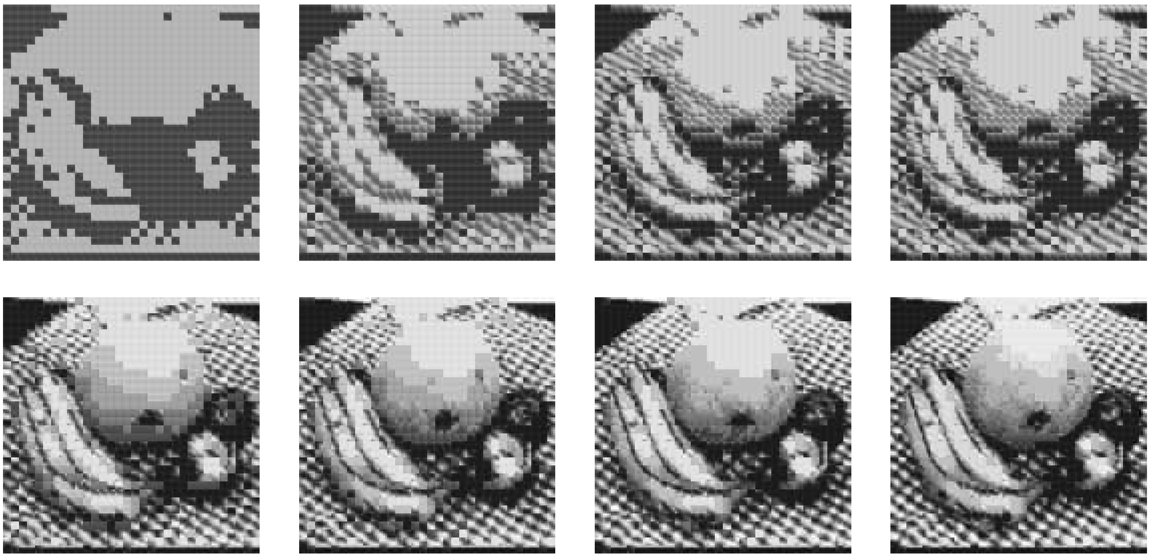

Figure 9: The reconstructed left image "Fruit" size 128x128 pixels in 1- 8 bits respectively based on QVQ technique using haar wavelet.

5 Conclusions

The disparity compensation can reduce stereo image size. The application of quadrant vector quantization that uses the advantage of coefficients symbol obtained by wavelet transform to make compact codebook. Our algorithm does not only reduce the number of codeword for training vector, but also gives serviceable quality of reconstructed image. Higher quality is obtained by smaller image experiment. In future work, we will experiment with another wavelet function. We plan to introduce overlapped block disparity compensation (OBDC) to reduce block artifact and comparing the results with other vector quantization algorithms.

6 References

[1] Mark S. Moellenhoff, Mark W. Maier, "Characteristics of disparity-compensated stereo image pair residuals", *Signal Processing Image Communication*, vol. 14, pp. 55-69 (1998).

[2] Michael G. Perkins, "Data Compression of Stereopairs", *IEEE Trans. Commun.,* vol.40, pp. 684-696, Apr. (1992).

[3] M. G. Strintzis and S. Malassiotis, "Object-based coding of stereoscopic and 3D image sequences: A review", *IEEE Signal Processing Mag.,* vol.16, pp. 14 -29, May. (1999).

[4] Nikolaos V. Boulgouris and Michael G. Strintzis, "A Family of Wavelet-Based Stereo Image Coders", *IEEE Trans. Circuits and Systems.,* vol.12, No.10, pp. 898-903, Oct. (2002).

[5] Qin Jiang, Joon Jae Lee and Monson H. Hayes, "A Wavelet Based Stereo Image Coding Algorithm", *ICASSP'99 Proceedings,* pp.3157-3160, (1999).