DISPARITY DEPENDENT SEGMENTATION BASED STEREO IMAGE CODING

Rahul Shukla[†], *Hayder Radha*[§] *and Martin Vetterli*[†]

[†]Audio-Visual Communications Laboratory
Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland
[§]Department of ECE, Michigan State University, East Lansing, USA
email: rahul.shukla@epfl.ch, radha@egr.msu.edu, martin.vetterli@epfl.ch

ABSTRACT

In this paper, we propose a novel rate-distortion (R-D) optimized disparity based coding scheme for stereo images. This new scheme efficiently integrates the coding of the disparity field with the residual image obtained via disparity estimation/compensation (DE/DC) in an R-D framework. The scheme first performs a quadtree decomposition of the target image and computes the disparity information along with the residual image for each node in the tree. An R-D based algorithm is then used for optimum bit allocation among the different quadtree nodes. The proposed scheme further jointly encodes the neighboring nodes with similar disparity information to attain higher coding gains. We present simulation results for the proposed scheme and compare these results with the performance of a fixed block size DE/DC based JPEG2000 stereo image coder. Our simulations show that the proposed scheme outperforms the fixed block size based disparity compensated JPEG2000 coder by more than 0.5 dB.

1. INTRODUCTION

It is well known that the binocular or disparity information extracted from stereo images plays a crucial role in several fields like computer vision, remote sensing/handling, tele-presence style video conferencing, tele-medicine and 3-D cinema. In particular, the majority of multi-view visual coding schemes are based on disparity-field estimation and compensation. Moreover, the recent increase in stereo visual applications translates into a growing demand for efficient methods for the transmission and storage of stereo image/video pairs. If the stereo images are compressed and transmitted independently using the standard coding schemes, then the required bandwidth would need to be doubled. However, due to the binocular dependency between the stereo images, they contain significant redundant information in the form of inter-frame redundancy [7]. Thus, by exploiting this binocular redundancy in addition to the intra-frame redundancy present in the two constituent images, we can achieve significant data compression without sacrificing the overall image quality.

Coding of stereoscopic images is generally based on exploiting the correlation between the left and right images. This is achieved by computing a disparity field between the stereo image pair [5, 7, 8]. The disparity field represents the amount of shift one needs to perform on the pixels within one image (target) to find the corresponding pixels in the other image (reference). A popular approach for computing the disparity field is to partition the target image into a set of blocks and perform a block matching algorithm to find the best match in the reference image. For coding applications, this approach is known as disparity estimation/disparity compensation (DE/DC) due to its resemblance to motion estimation/motion compensation (ME/MC) methods, which are popular for video coding. The key concept of stereo image compression based on DE/DC is to use one of the images in the stereo pair as a reference and to predict the other image (the target).

Disparity compensation based prediction for stereo images was first introduced by Lukacs [5]. In [7], Perkins proposed the conditional coder/decoder structure for the stereo image coding and analyzed the R-D behavior of this structure. Aydinoglu *et al.* presented a coding scheme which combines disparity compensation and transform coding of the residual image in the single framework using an adaptive incomplete transform [1]. Hierarchical block matching algorithms have been used to generate a more homogeneous disparity fields that lead to better coding efficiency [10, 15].

In [2], a novel scheme is presented for the efficient exploitation of the zerotree algorithm in stereo image coding applications. Disparity field estimation based on low level features, such as edges, and on model/object based methods have also been used [13]. In [15], feature/object based DE/DC scheme is proposed. This scheme determines a set of objects/features in both images and seeks correspondence between the two sets. Jiang [3] proposed a hybrid approach

This work was supported by the Swiss National Science Foundation grant 20-63664.00.

between block and object based schemes for the efficient coding of stereo pairs. In [18], overlapped block disparity compensation with adaptive search windows is presented. In [6], Moellenhoff *et al.* analyzed the properties of disparity compensated residual images and proposed transform domain coding methods for improved coding of residual images. Other efforts on stereo coding have focused on optimal rate-distortion based algorithms for block dependent quantization [17] and motion/disparity field estimation [16].

Our goal is to develop a compression algorithm, which performs the disparity dependent segmentation of stereo images in the R-D framework. We employ a generalized quadtree decomposition based DE/DC. This scheme further performs the joint coding of the disparity information and the residual image obtained by the disparity compensation in the R-D optimal sense.

The rest of the paper is organized as follows: In Section 2, we describe the proposed disparity dependent segmentation based coding scheme for stereo images. We then present some simulation results in Section 3. Finally, we conclude with a discussion of further research directions in Section 4.

2. DISPARITY DEPENDENT SEGMENTATION BASED STEREO IMAGE CODING ALGORITHM

The generic structure of a stereo image coding scheme is as follows: First encode one of the images of the stereo pair as reference, then estimate the disparity information between blocks in the target and encoded reference images and code the disparity information and the residual image obtained by subtracting the disparity compensated target image from the original image. The disparity compensated/predicted target image is obtained by using the encoded reference image and the disparity map. Thus, our scheme employs the closed loop methodology for obtaining the predicted target image.

The main novelty of our stereo image coding scheme is that it performs the disparity dependent quadtree segmentation in the R-D framework.[1] This disparity dependent segmentation based stereo image coding scheme can be described as follows:

Step 1: Disparity estimation and compensation

1. Segment the target image using the quadtree decomposition up to a tree depth J.
2. Compute the disparity information for each node by using an intensity based block matching approach.
3. Predict the node using the associated disparity information and encoded reference image.
4. Compute the residual image block by subtracting the predicted target image block from the original image block.
5. Generate the R-D curve for the residual block associated with each node using a standard coding scheme like JPEG2000 [14].

Step 2: The Lagrangian cost $(L(\lambda) = D + \lambda R)$ **based pruning**

6. For the given operating slope λ, R-D optimal pruning criterion is as follows: Prune the children if the sum of the Lagrangian costs of the children is greater than or equal to the Lagrangian cost of the parent. This parent-children pruning criterion is used recursively to do fast pruning from the full tree depth towards the root to find the optimal subtree for a given λ [9]. This scheme provides a pruned tree. However, this independent pruning scheme fails to join neighboring blocks with similar disparity information, if they have different parents. Thus, this coding scheme fails to exploit the complete dependency among neighbors. For correcting this suboptimality, we introduce the following neighbor joining step proposed in [11, 12].

Step 3: Neighbor joining

7. Given the pruned tree obtained from Step 2, the neighbor joint coding is performed on the leaves of the pruned tree in the R-D optimal manner. That means, the two neighbors will be joined if the sum of the Lagrangian costs of the neighbors is greater than or equal to the Lagrangian cost of the joint block. This scheme essentially ensures that the neighbors with similar disparity information are jointly coded for further improving the R-D performance. Thus, this leads to the disparity dependent segmentation of the target image.
8. By summing up the allocated rates to all the joint blocks along with the costs of the segmentation map and disparity map will provide the overall bit-rate $R^*(\lambda)$. Similarly, by summing up the associated distortions of all the joint blocks will give the net distortion $D^*(\lambda)$. Since we have employed the R-D optimization, $D^*(\lambda)$ represents the minimum distortion for the bit-rate constraint $R^*(\lambda)$.

Step 4: Search for the desired R-D operating slope

9. The value for λ is determined iteratively until the bit-rate constraint R_0 is met as closely as possible. We employ the fast bisection search algorithm given in [9].

Clearly, the above described algorithm provides the R-D optimal representation of the target image for a given bit budget and encoded reference image. Due to neighbor joining (Step 3), it is obvious that the proposed scheme results into the non regular partitioning of the target image (for example, see Figure 3). It is important to note that the high quality encoded reference image tends to allow for more efficient encoding of the residual image due to the good quality of the disparity predicted target image.

Similar to the encoder, the structure of the decoder is as follows: 1) Reconstruct the reference and residual images. 2) Generate the predicted target image using the disparity information and reconstructed reference image. 3) By summing the predicted target image and residual image, we obtain the desired target image.

[1] Mean squared error (MSE) is used as a distortion measure.

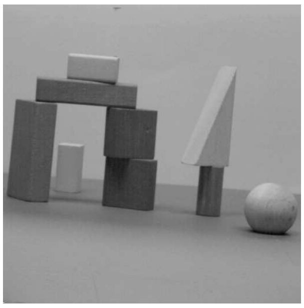
3. SIMULATION RESULTS

Simulations are performed on the Arch stereo image pair shown in Figure 1.[2] This stereo pair has the disparity predominantly along the horizontal direction, which is usually the case in stereo image coding scenario due to the parallel axis geometry. Therefore, for DE/DC, we utilize the search window of size 64 in the horizontal direction whereas search in the vertical direction is confined only to ± 2 pixels.

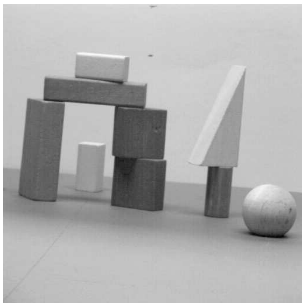
Figure 2 shows the disparity field obtained by the quadtree based variable block size decomposition of the target image. As expected, the disparity map is smooth everywhere except along the object boundaries and in the occluded regions.

A reconstructed target image along with the disparity dependent segmentation map is presented in Figure 3. It is evident from Figure 3 that the algorithm correctly performs finer segmentation along the edges of objects to capture the depth discontinuity precisely, whereas smooth regions with uniform disparity are represented by larger blocks. This results into smaller disparity information overhead bits for smooth regions. This figure also demonstrates that the proposed coding scheme is capable of achieving high compression ratios without sacrificing the reconstructed image quality which we are measuring in terms of PSNR. These simulations also indicate that our R-D optimized algorithm automatically takes care of occluded regions either by performing finer segmentation in those regions or by allocating more bits to corresponding residual blocks.

Figure 4 compares the R-D performance of our disparity dependent segmentation based coding scheme with the independent JPEG2000 coder and fixed block size (16×16) based disparity compensated JPEG2000 stereo image coder. This R-D performance comparison clearly shows the superiority of the proposed algorithm over the other two algorithms.



(a) Reference image.



(b) Target image.

Fig. 1. Original Arch stereo pair.

[2]These stereo images are obtained from http://vasc.ri.cmu.edu//idb/html/stereo/arch/index.html.

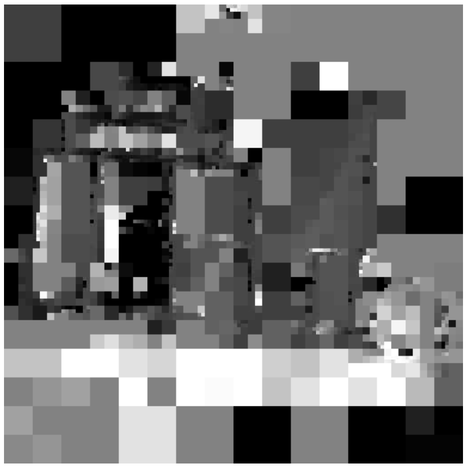


Fig. 2. Quadtree based disparity map.

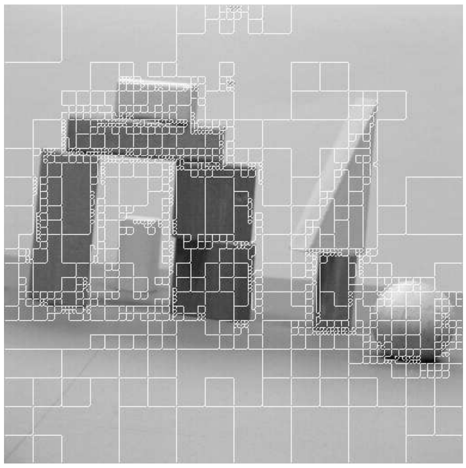


Fig. 3. Reconstructed target image along with the segmentation map obtained by the disparity dependent segmentation based coding scheme (PSNR=41.4 dB, Bit-rate=0.116 bpp).

4. DISCUSSION AND FURTHER WORK

We have presented an R-D optimized disparity dependent segmentation based stereo image coding scheme. This new scheme performs the joint coding of disparity information and the residual image to achieve the improved R-D performance. Experimental results shown in Figure 4 also confirm that the proposed scheme outperforms the disparity compensated JPEG2000 stereo coder by about 0.5 dB. Since our scheme employs a quadtree based decomposition, this scheme is computationally efficient as well. Our ongoing research effort is to extend this stereo image coding scheme to multi-view image and stereo video coding scenarios. We would also like to investigate feature/object based DE/DC scheme in the framework of the proposed algorithm.

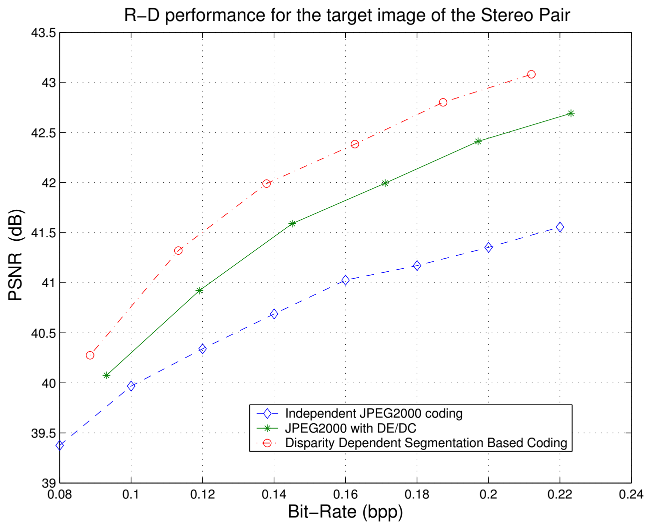


Fig. 4. R-D performance comparison of different coding schemes for the target image of the Arch stereo pair shown in Figure 1.

Another interesting problem setup is the distributed coding scenario for stereo images. For instance, consider an electronic surveillance system where the two cameras are viewing the same scene. In this setup, the communication between the two encoders is not allowed due to some constraints such as the power constraint. We are planning to develop a distributed stereo image coding algorithm for the cases where the communication cost between the encoders is much higher than the computation cost.

5. REFERENCES

[1] H. Aydinoglu and H. Hayes, "Stereo image coding: A projection approach," *IEEE Trans. Image Proc.*, vol. 7, pp. 506-516, Apr. 1998.

[2] N. V. Boulgouris and M. G. Strintzis, "A family of wavelet-based stereo image coders," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 12, no. 10, pp. 898-903, Oct. 2002.

[3] J. Jiang and E. A. Edirisinghe, "A hybrid scheme for low bit-rate coding of stereo images," *IEEE Trans. Image Proc.*, vol. 11, no. 2, pp. 123-134, Feb. 2002.

[4] F. Labonte, C. T. Le Dinh, J. Faubert, and P. Cohen, "Spatiotemporal spectral coding of stereo image sequences," *IEEE Trans. Circuits, Syst., Video Tech.*, vol. 9, pp. 144-155, Feb. 1999.

[5] M. E. Lukacs, "Predictive coding of multi-viewpoint image sets," in *Proc. ICASSP*, Oct. 1986, pp. 521-524.

[6] M. S. Moellenhoff and M.W. Maier, "Transform coding of stereo image residuals," *IEEE Trans. Image Proc.*, vol. 7, pp. 804-812, June 1998.

[7] M. G. Perkins, "Data compression of stereo pairs," *IEEE Trans. Comm.*, vol. 40, pp. 684-696, Apr. 1992.

[8] A. Puri, R. V. Kollarits, and B. G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4," *Jour. Signal Proc. Image Comm.*, vol. 10, pp. 201-234, 1997.

[9] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate distortion sense," *IEEE Trans. Image Proc.*, vol. 2, no. 2, pp. 160-175, April 1993.

[10] S. Sethuramam, M. W. Siegel, and A. G. Jordan, "A multi-resolution framework for stereoscopic image sequence compression," in *Proc. ICIP-95*, vol. 2, pp. 361-365.

[11] R. Shukla, P.L. Dragotti, M.N. Do and M. Vetterli, "Improved quadtree algorithm based on joint coding for piecewise smooth image compression," in *Proc. IEEE Int. Conf. on Multimedia*, Switzerland, Aug. 2002.

[12] R. Shukla, P.L. Dragotti, M.N. Do and M. Vetterli, "Rate-distortion optimized tree structured compression algorithms for piecewise smooth images ," submitted to *IEEE Trans. Image Proc.*, Dec. 2002.

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

[14] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, Dordrecht, 2001.

[15] D. Tzovaras, N. Grammalidis, and M. G. Strintzis, "Object-based coding of stereo image sequences using joint 3D motion/disparity compensation," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 7, pp. 312-327, Apr. 1997.

[16] D. Tzovaras and M. G. Strintzis, "Motion and disparity field estimation using rate-distortion optimization," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 8, pp. 171-180, Apr. 1998.

[17] W. Woo and A. Ortega, "Optimal blockwise dependent quantization for stereo image coding," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 9, pp. 861-867, Sept. 1999.

[18] W. Woo and A. Ortega, "Overlapped block disparity compensation with adaptive windows for stereo image coding," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 10, pp. 194-200, Mar. 2000.