To be published in SPIE Vol. 2668 - Digital video compression - Algorithms and Technologies - 1996

SEGMENTATION BASED CODING OF STEREOSCOPIC IMAGE SEQUENCES

Sriram Sethuraman, M. W. Siegel, Angel G. Jordan

Department of Electrical and Computer Engineering
The Robotics Institute / School of Computer Science
Carnegie Mellon University, Pittsburgh, PA 15213

ABSTRACT

A binocular disparity based segmentation scheme to compactly represent one image of a stereoscopic image pair given the other image was proposed earlier by us. That scheme adapted the *excess bitcount*, needed to code the additional image, to the binocular disparity detail present in the image pair. This paper addresses the issue of extending such a segmentation in the *temporal dimension* to achieve efficient stereoscopic *sequence* compression. The easiest conceivable temporal extension would be to code one of the sequences using an MPEG-type scheme while the frames of the other stream are coded based on the segmentation. However such independent compression of one of the streams fails to take advantage of the segmentation or the additional disparity information available. To achieve better compression by exploiting this additional information, we propose the following scheme. Each frame in one of the streams is segmented based on disparity. An MPEG-type frame structure is used for motion compensated prediction of the segments in this segmented stream. The corresponding segments in the other stream are encoded by reversing the disparity-map obtained during the segmentation. Areas without correspondence in this stream, arising from binocular occlusions and disparity estimation errors, are filled in using a disparity-map based predictive error concealment method. Over a test set of several different stereoscopic image sequences, high perceived stereoscopic image qualities were achieved at an excess bandwidth that is roughly 40% above that of a highly compressed monoscopic sequence. Stereo perception can be achieved at significantly smaller excess bandwidths, albeit with a perceivable loss in the image quality.

Keywords: stereoscopic sequence compression, disparity based segmentation, quadtrees, multiresolution, error concealment

1. INTRODUCTION

Stereoscopic sequences can be compressed much more efficiently than the independent compression of its two image streams by exploiting, in addition to the spatial and temporal correlations that are exploited by sequence compression schemes, the high cross-stream correlations present between the streams. Several researchers have suggested such joint coding of the two streams(6,8,10,11,13). We have progressively addressed the specific problem of coding stereoscopic image sequences at low excess bandwidths over the bandwidth required for monoscopic sequences, without unduly sacrificing the perceived stereoscopic image quality(5, 3, 1). The conclusion from these studies was that we can achieve significant reduction in overall bit rate on the average by shifting to content-based adaptive coding strategies.

A disparity-based segmentation and the subsequent transmission of these segment disparities was proposed by us[1] to encode one image of a stereoscopic image pair at a very low coding overhead given the other image. This scheme adapts the overhead to the disparity detail present in a given stereoscopic image pair, unlike a fixed block size based scheme. Also, by segmenting at object boundaries, it reduces the number of spurious matches while preserving disparity discontinuities.

This paper treats the system level integration needed to employ this segmentation in the compression of stereoscopic sequences. Tracking the segments across frames leaves unpredicted regions in the estimated frames. Thus the most natural schemes may not provide the most visually pleasing results at low bit rates. Several possible temporal extensions are considered, and the particular scheme that provides a good quality stereoscopic stream at high compression rates, albeit at a higher computational expense, is presented here.

The paper is arranged as follows. Section 2 provides a brief review of the disparity-based segmentation scheme. Section 3 discusses three different means of stereoscopic sequence compression incorporating this segmentation, and provides the justifications for choosing a particular scheme. Section 4 elaborates on the specific scheme chosen. Section 5 presents a performance evaluation in terms of the excess bandwidth needed to transmit the additional sequence and the corresponding signal-to-noise ratios (SNRs) that are obtained for three different stereoscopic sequences. Section 6 summarizes the paper and outlines a few possible directions for further work.

2. DISPARITY BASED SEGMENTATION - REVIEW

Typical stereoscopic images contain large regions of almost constant binocular disparity arising from the scene backgrounds and large objects at a fixed depth. Fixed block size (FBS) based disparity estimation schemes divide these regions into smaller blocks, thus requiring more block disparities to be coded than is necessary. Also, matching small featureless blocks leads to spurious matches which affect the smoothness of the estimated disparity map and render the differential predictive coding of the block disparities ineffective. Also, block based disparity estimation fails for blocks containing portions of two objects at different depths. To overcome these drawbacks, and to achieve very low overhead for coding one image of the stereoscopic pair given the other, a multiresolution and disparity based segmentation (DBS) scheme was proposed by us[1]. For the sake of continuity, a brief overview of the scheme (with the few modifications since [1]) is presented here.

The stereoscopic image pair is subjected to multiresolutional decomposition to get progressively lower resolution images. The segmentation proceeds from the coarsest resolution level to the original resolution level. At each resolution level, a block is recursively split horizontally and / or vertically depending on the disparity estimated for each of the four subblocks. A block is not split if all four of its subblocks have the same disparity or if the size of the block is below a preset value. The splitting locations are obtained using a simple ‘dominant edge selection’ algorithm (described in [1]) from the block’s intensity values, thus making the segments align with object boundaries (which usually occur at an intensity discontinuity). Thus, this scheme adapts the size of the segments according to the disparity detail present in the stereoscopic pair. This significantly reduces the number of bits needed to represent one image of the stereoscopic pair given the other, even after taking into account the overhead for the representation of the segmentation.

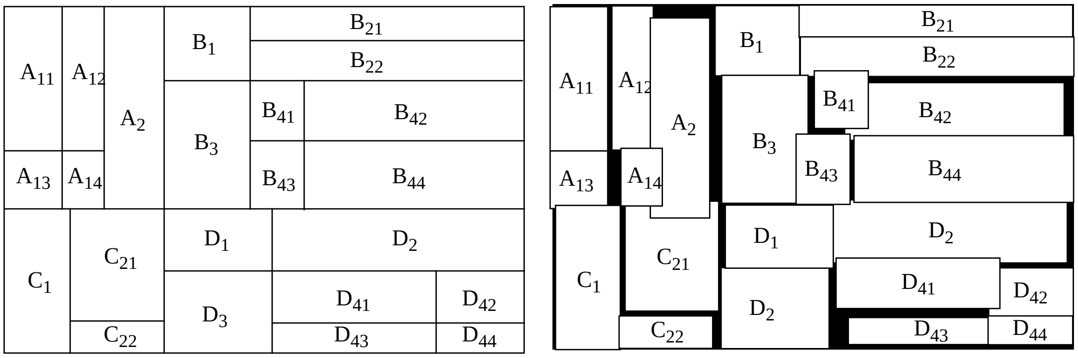
3. SEGMENTATION BASED STEREOSCOPIC SEQUENCE COMPRESSION

The segmentation scheme described in the previous section can be incorporated in a stereoscopic sequence compression framework in several different ways. Our goal is to choose a system level scheme that results in an overall minimum bit rate for the jointly coded stereoscopic stream at reasonably high stereoscopic image qualities, while maintaining a moderate computational complexity at the encoder and low computational complexity at the receiver.

3.1 Reversal of prediction direction:

Before considering the different options, we introduce the concept of ‘reversal of prediction direction’ and its impact on the coding of the predicted image, so that the comparisons later on can be better understood. Typical motion estimation schemes partition the image to be coded into non-overlapping blocks and then find the best match for each segment in the reference image. Thus even for occluded regions, a reasonably close match can be obtained by making the segment sizes smaller. Now let us consider the reverse case in which the reference image (A) is partitioned into non-overlapping segments, and the image (B) is estimated by finding the best match in B for each segment in A. This situation arises if the segments are to be *tracked* in the temporal dimension, or in the other view, or both; we refer to it as the ‘reversal of prediction direction’.

The impact of such reversal in the direction of prediction is illustrated in Fig.1. As objects within the scene undergo displacement, new regions may be exposed and currently exposed regions may get occluded. If a segment in A is occluded partially in B, then the best match for that segment can occur at the correct location or a spurious match may be generated, depending on the extent of occlusion and the chance existence of other good matches. When the match occurs at the correct location, the occluded region has two candidate matches - one corresponding to the occluded region and the other corresponding to the occluding region. For example, in frame B, a portion of segment B_{41} occludes a portion of segment B_{22}. The common region between these two segments thus has two possible candidate matches. When a spurious match occurs, the corresponding segment leaves behind an unfilled region and also adds itself as a candidate estimate for the false match location.



Segmented frame - A (non-overlapping partition)

Frame being estimated - B

Corresponding segments in B overlap due to motion occlusions & motion estimation errors
New regions (shown black) that were occluded in A are exposed in B

Fig.1: Impact of the reversal of prediction direction

To code the predicted image, we need to (a) select the correct match when there are multiple matches and (b) fill in the pixel values in the exposed regions. The former can be resolved if a reliable depth (or disparity) map of the pixels in A are available. The latter requires encoding the exposed region either through estimation or intracoding. Typically the number of such exposed regions is of the order of the number of segments. Hence estimation-based filling in may offset any gain achieved through segmentation. Also a fixed radix DCT based intracoding is not possible due to the irregular shape of the exposed regions. Thus encoding these regions poses a significant overhead compared to the case in which the prediction direction is not reversed. However most of the exposed regions are only a few pixels wide at least in one direction and thus a properly implemented interpolation scheme can fill in most of the regions, albeit at a higher computational expense at the decoder. The regions which are too wide to fill in can then be estimated or intracoded.

3.2 Scheme-1:

The simplest possible approach to incorporate the DBS scheme in stereoscopic sequence coding would be to encode one image sequence independent of the other stream. Each frame of the second stream is segmented based on the disparity estimated with respect to the independently coded stream and the segment disparities are used to represent this second stream. Thus the segmentation information is used in coding only one stream. Since the segmentation is repeated for every frame, this scheme is compute intensive. To warrant this additional computational burden, the segmentation overhead bits combined with the segment disparity coding bits should be fewer than the disparity coding bits needed in a FBS based scheme. As the prediction direction is not reversed in this scheme, the issues mentioned in section 3.1 do not arise. The independently coded stream can also be efficiently coded by employing motion-based segmentation (MBS). However, this would result in significantly increased complexity and it also doubles the segmentation representation coding overhead.

3.3 Scheme-2:

An approach that would require less frequent segmentation is one in which the DBS is performed at a certain frame, and the segments are tracked over the next several frames in both the left and right streams until the scene content changes significantly. This scheme is more natural in the sense that only the newly exposed regions need to be coded in addition to specifying the offsets for the tracked segments. It also offers reduced complexity and the advantage that the receiver can use the same set of segments in the reference frame to decode all intercoded frames. However, the prediction direction is reversed for all the intercoded frames. The effort of ‘filling in the holes’ can offset all the other advantages and the quality does not degrade gracefully at lower bit rates. Also, different regions within a segment obtained using DBS can undergo different displacements from frame to frame. Hence the segments from DBS would have to be further decomposed as necessary for better motion compensation.

3.4 Scheme-3:

Since the more natural scheme-2 does not perform well from the coding point of view, we modify scheme 1 so that both the streams are efficiently coded based on segmentation and only one segmentation is performed per stereoscopic pair of frames. Each frame in one of the video streams (main stream) is segmented based on disparity with respect to the other stream (auxiliary stream). The best matches for the segments are obtained from the reference frames for that video stream. The segments are further decomposed as necessary to account for different displacements within the same segment. The direction of the disparity maps obtained during the segmentation is *reversed* to estimate the frames in the auxiliary stream. A schematic diagram is shown in Fig.2.The ‘holes’ arising from the reversal of the prediction direction are concealed by spatial interpolation while making use of the disparity to avoid interpolating across regions at different depths. Concealment in this case is easier because of the structure introduced by the parallel-axes binocular imaging geometry (1, 4). Concealment in scheme-2 is much harder because multiple objects within the scene can move in arbitrary directions.The disparity map is also used to resolve among multiple matches. Since segmentation is performed for every frame, the computational burden for this scheme is higher than that of scheme-2. Implementation specifics of scheme-3 are discussed in detail in the next section.

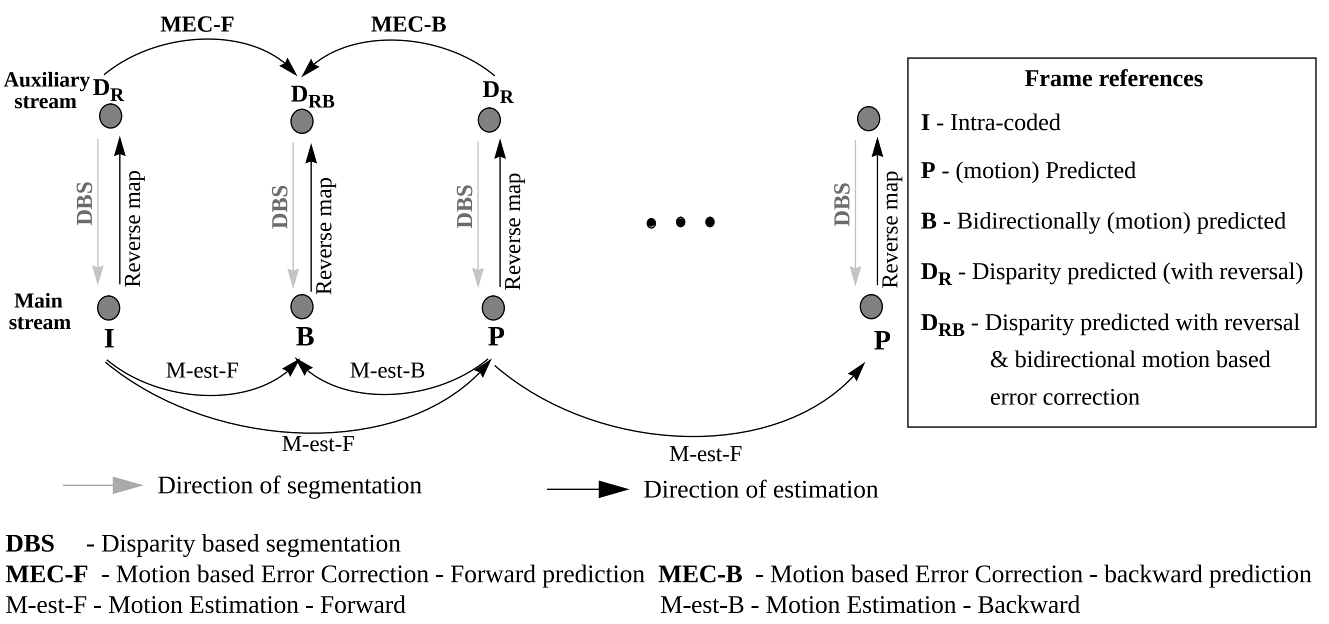


Fig. 2: Schematic diagram of stereoscopic sequence compression scheme-3

4. IMPLEMENTATION DETAILS

The sequence compression approach in scheme-3 meets the objectives set forth in the introduction to section 3 and hence is chosen for implementation. This section explains the scheme (shown in Fig.2) in more detail. It also addresses the issues of error concealment, and residual coding of the Y, U and V components.

4.1 Structure of the sequence compression scheme:

The implementation considers a frame structure similar to that of the MPEG standards. The structure and the dependencies between the frames are illustrated in Fig.2. Two new frame types, namely the D_R and D_{RB} frames, are considered in addition to the intra-coded (I), predicted (P) and bidirectionally predicted (B) frames in order to represent the frames in the additional stream in a stereoscopic sequence. The stream consisting of the I, P and B frames will be referred to hereafter as the *main stream*, because this stream is always coded at a high quality to maintain compatibility with existing monoscopic transmissions. The other stream will be referred to as the *auxiliary stream*, as its quality can be adjusted depending on the available excess bandwidth.

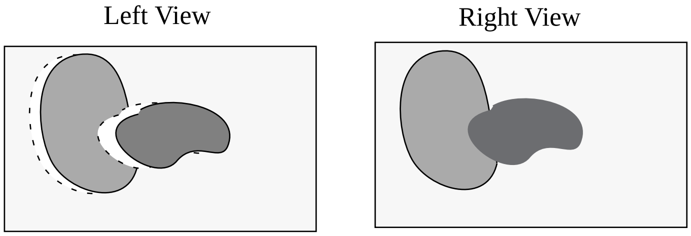
An I-frame is segmented based on disparity. The non-overlapping segments can be efficiently intra-coded by finding a 2-D polynomial fit to the intensity profile within the segment(2) and coding the resulting coefficients, the segmentation overhead and the residuals. The segment disparities computed during the segmentation are reversed to estimate the corresponding frame in the auxiliary stream (D_R-frame). The disparity estimates are also used to resolve among multiple candidates in overlapping regions. Error concealment (discussed later in 4.2) and residual coding are carried out to boost the quality of the D_R frame.

A P-frame is segmented based on disparity and a motion estimation is performed to obtain the best match for the segments in the previous I or P frame. The segmentation overhead, motion vectors and the residuals are coded. The corresponding frame in the auxiliary stream is also a D_R frame, and is processed in the same way as the D_R frame corresponding to an I-frame.

The B-frames are segmented based on disparity. The segments are bidirectionally motion predicted with respect to the past and future reference (I or P) frames. The segmentation overhead, the forward/backward motion vectors and the residuals are coded. The corresponding frames in the auxiliary stream (D_{RB}-frames) are obtained by reversing the disparity map obtained from the DBS and then concealing the ‘holes’ (described later). The regions in a D_{RB} frame that are occluded in the corresponding B frame (and hence could not be predicted from that B-frame) are bidirectionally motion predicted from the past and future D_R frames with fixed block sizes. Since the D_{RB} frames are not used by any other frame as a reference, only visually objectionable residuals are coded.

4.2 Error concealment following disparity map reversal:

Given that the stereoscopic pair of frames are captured at the same time, the disparity map depends only on the depths of the different objects at that instant and the fixed binocular camera geometry. Thus, occlusions due to binocular parallax (regions that are visible in one view and not in the other) are more structured, unlike motion-based occlusions which depend on the displacements of the different objects in the scene. In particular, when the parallel axes binocular imaging geometry(1, 4) is used, the disparity has only a horizontal component. Also, assumptions of translation-only displacement are more likely to be valid for stereoscopic cameras with fixed focus wide-angle lenses. In contrast, individual objects in the scene can have rotation, scale and shear components temporally. A simple error concealment scheme for a binocular occlusion would be to assume that this region belongs to an object that is at a greater depth than the occluding object. A simple case is illustrated in Fig.3. The object to the right is in front of the object to the left. The areas that will be exposed when the left view is estimated from the right are shown in white. Thus, operating along each scan line (since only the horizontal component of disparity is



The area between the dotted and the dark lines in the left view shows the area that is occluded in the right view. Most of the exposed area can be filled in by extending the value to the left of the occluded area.

Fig. 3a Fig. 3b Fig. 3c

Fig.3: Concealment of perspective-based occlusions

(a) A D_{BR} frame estimated by reversal of disparity map (regions without estimate appear as black patches)

(b) Frame in (a) after the described concealment (c) The original frame (for comparison)

present), we can fill in the exposed regions by extending the intensity values of a nearby object. In most typical scenes, this filled in value is close to the actual value for a large fraction of the exposed regions. This type of extension also preserves the continuity in the intensity profile and thus conceals the errors very well. A D_{BR} frame estimated by reversal of the disparity map is shown in Fig. 3a. The same frame after concealment is shown in Fig. 3b. The original frame is shown in Fig. 3c for comparison.

4.3 Residual coding:

Motion or disparity compensated prediction provides a reasonable estimation for a fairly large fraction of a frame. However, such a predicted frame can contain visually unpleasant artifacts due to the failure of the assumption of fixed displacement or disparity over an entire block (or) the failure of the concealment procedure outlined in section 4.2. Since all frames other than the D_{RB} frames are used as references in predicting other frames, these frames need to be coded at a quality commensurate with the extent of their usage as reference frames. Hence a residual coding scheme with a smooth rate control is essential for final visual quality. The residual coder used in our simulations is a quadtree based residual vector quantizer (VQ). Each 16x16 block is recursively split into 4 subblocks based on the mean squared error (MSE) of the block. An 8x8 block is not divided if a good match can be found in one of the reference frames (as was discussed in section 4.1). Blocks of size 4x4 (2x2) are not split if a good match for the residuals is obtained within a fixed 16-dimensional (4-dimensional) vector codebook. The single pixels are coded using a scalar quantizer (SQ). The encoded quadtree structure and the entropy coded VQ and SQ indices constitute the residual coding overhead.

4.4 Chrominance components processing:

The 4:2:0 format (horizontal and vertical subsampling of U and V by a factor of 2) recommended by the MPEG committee is used as the input to the sequence compression scheme. A half-band low pass filtering is carried out before subsampling to minimize aliasing. After computing the motion / disparity of a segment using only the luminance signal in the segmentation and estimation processes, the estimates are halved to obtain the estimates for the corresponding U and V segments, which are at half the resolution. Any fractional pixel accuracy is ignored in obtaining these estimates. The error concealment in 4.2 works equally well for the Y, and U, V components.

Since the chrominance components are smooth, the mean value of the residuals within a block is used to code a block corresponding to the 4x4 and 2x2 block sizes in the residual quadtree described in the last section. Hence the residual coding overhead for the U and V components is very small compared to that of the Y component.

5. PERFORMANCE EVALUATION

We report here the compression ratios and corresponding SNRs that result from employing our compression scheme on three stereoscopic sequences that were recorded using a fixed focus, field sequential stereoscopic camera (with 9.5mm lenses). We name these sequences as *booksale*, *buggy* and *crowd* sequences† respectively to indicate the type of activity within each sequence. The *booksale* sequence involves a panning camera, has small displacements of objects within the scene and conveys a good sense of depth. The *buggy* sequence has a very high motion content (both object as well as camera motions) and a moderate depth range. The *crowd* sequence has very little camera motion, random object displacements within the scene, and a good depth range with objects at several depths that occlude each other due to the displacements. Thus, these sequences constitute a diverse set of sequences for testing the performance of the compression scheme.

Without loss of generality, the distance between I and P (or P and P) frames is chosen as M=3 frames. Though the separation between I-frames in an actual implementation should depend on the outputs of a scene change detector, we assume a fixed separation of N=16 frames for simplicity. We define the compression ratio and the excess bandwidth factor as follows:

$$\text{compression ratio (CR)} = \frac{\text{bits per pixel for the actual sequence}}{\text{average bits per pixel for the compressed sequence (av. bpp)}}$$

$$\text{excess bandwidth factor} = \frac{\text{average bits per pixel for the auxiliary stream}}{\text{average bits per pixel for the main stream}}$$

The original bpp in the case of the 4:2:0 YUV format is 12 bits. The bit rates reported are after suitable entropy coding of the motion / disparity estimates and the VQ / SQ indices. The segmentation overhead is coded by specifying the offset of the horizontal and vertical split locations with respect to the height and width of a block. The residual quadtree structure is coded with 1 bit per split. Though run length coding of these bits can further reduce the bit rate, it has not been implemented.

In line with our objectives stated in the beginning of section 3, the main stream is coded at a higher bit rate to preserve the quality of one of the streams. This bit rate is kept constant though out the experiments. The bit rate for the auxiliary stream was varied smoothly in the following fashion. The bit rate resulting from coding only the segment disparities (reversal of disparity map followed by concealment) serves as the datum (or) the bare minimum needed to code the auxiliary stream. The second level of improvement in quality at low bit rates is achieved from the motion estimation based filling in of the occlusions and errors. The third level of improvement in quality is achieved through the quadtree and VQ / SQ based residual coding. A smooth control over the bit rate in the last two cases is achieved by setting thresholds for the mean squared error and

---

†. We can make these sequences available to other research groups working on stereoscopic image sequence compression, to enable comparisons with other existing schemes.

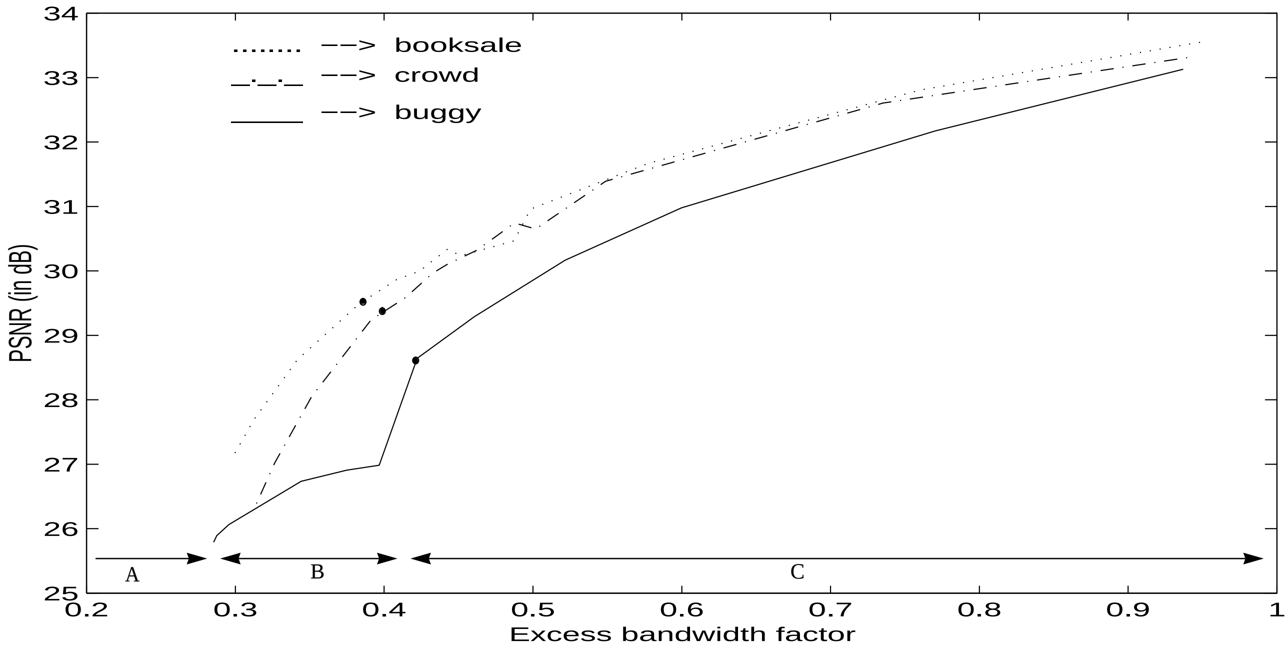


Fig. 4: Average PSNR Vs. Excess bandwidth factor for the D_{BR} frames

- The points above which the perceived stereoscopic quality is good (the corresponding data appear in Table I)

REGION A: Only the segment disparities are coded.

REGION B: Bidirectional motion prediction to fill in occlusions is added.

REGION C: Residual coding is added to the above two.

For the *buggy* sequence, the curve shows no significant PSNR improvement in region B due to the high motion content.

the *significant error count* (presence of errors greater than a threshold). The *excess bandwidth factor*s and the resulting average peak signal to noise ratios (PSNR) for the D_{BR} frames are plotted in Fig. 4. The excess bandwidth factor tends to one as the PSNR for the D_{BR} frames tends to the PSNR of its reference frames. Subjectively, the perceived stereoscopic image quality is quite good above excess bandwidth factors near 0.4 (the dark dots in Fig. 4). This region of the curve corresponds to the D_{BR} frames after bidirectional motion prediction to fill in the occlusions and errors. However in the curve corresponding to the *buggy* sequence, owing to the failure of translation only displacement assumptions, this stage could not boost the PSNR and residual coding is required to achieve good viewing quality.

TABLE 1: Compression ratios and the associated average PSNRs

Sequence name	Main stream						Auxiliary stream				Compression ratio (CR)			Excess bandwidth factor
	I		B		P		D_R		D_{RB}					
	PSNR	bpp	PSNR	bpp	PSNR	bpp	PSNR	bpp	PSNR	bpp	Main stream	Aux. stream	Overall	
Booksale	36.19	1.15	34.11	0.218	34.14	0.246	32.6	0.196	29.48	0.08	42.25	110.1	61.07	0.383
Buggy	36.15	1.19	33.80	0.27	33.97	0.251	32.64	0.196	28.64	0.119	36.36	86.33	51.17	0.421
Crowd	35.50	1.18	33.50	0.237	33.70	0.296	32.16	0.217	29.29	0.088	38.83	99.17	55.81	0.391

Table I lists the average bpp for the different frame types and the corresponding average PSNRs for the luminance components of the three sequences. The data for the D_{BR} frames corresponds to the point above which the perceived stereoscopic quality was quite good.

The human visual system perceives a sharp stereoscopic image of good quality even when one of the streams is coded at a lower quality, as long as no visually distracting artifacts are present. Thus, the stereoscopic sequence can be transmitted for about 1.4 times the monoscopic transmission bandwidth. Since the main stream is also coded based on segmentation, even lower bit rates than the ones shown in Table I provide quite acceptable results. Thus, the joint bpp for the stereoscopic sequence can be made closer to the average bpp for a single sequence coded using fixed block based schemes without any noticeable distortions in the main stream as well.

6. CONCLUSIONS & FUTURE WORK

We have presented in this paper a unified framework for stereoscopic image sequence compression that achieves scene-adaptive bit rates by incorporating the binocular disparity-based segmentation at the heart of the scheme. A bit allocation scheme that can smoothly vary the excess bandwidth has been implemented. Experimental results on three representative sequences were presented. These results show that low excess bandwidths can be achieved by incorporating the disparity-based segmentation to estimate the auxiliary stream from the main stream and by coding the auxiliary stream at a slightly lower quality than the main stream. The slight loss in the quality of the auxiliary stream is not noticeable when the decoded sequences are viewed stereoscopically. Segmentation-based coding of the main stream results in lower bit rates for the main stream also.

We are currently quantifying the advantages of segmentation based coding of the main stream compared to a fixed block size based scheme. We also intend to explore the extensibility of this scheme to a symmetric extension of the binocular imaging setup, in which three cameras are used. The middle (cyclopean[14]) camera can be used to generate a high definition monoscopic stream, whereas the left and right (possibly low definition) cameras can be used to generate the left and right disparity streams. The symmetric treatment is intended to relieve any eyestrain that may be caused due to prolonged viewing of the two sequences at different image qualities.

7. ACKNOWLEDGEMENTS

This research was supported by the Advanced Research Projects Agency under ARPA Grant No.MDA 972-92-J-1010.

8. REFERENCES

1. S.Sethuraman, M.W. Siegel, A.G. Jordan, ‘A multiresolutional region based segmentation scheme for stereoscopic image compression’, Proc. of the IS&T/SPIE Symp. on Electronic Imaging, Digital Video Compression-Algorithms and Technologies 1995, Vol.2419, 1995.
2. X. Wu and Y. Fang, ‘A segmentation-based predictive multiresolution image coder’, IEEE Transactions on Image processing, Vol.4, pp.34-47, January 1995.
3. S.Sethuraman, M.W. Siegel, A.G. Jordan, ‘A multiresolution framework for stereoscopic image sequence compression’, Proc. of ICIP ‘94, Vol. II, pp.361-365, IEEE Computer Society Press, 1994.
4. V.S. Grinberg, G. Podnar, M.W. Siegel, ‘Geometry of binocular imaging’, Proc. of the IS&T/SPIE Symp. on Electronic Imaging, Stereoscopic Displays and applications, Vol.2177, 1994.
5. S.Sethuraman, M.W. Siegel, A.G. Jordan, ‘Multiresolution based hierarchical disparity estimation for stereo image pair compression’, Proc. of the symposium on Application of subbands and wavelets, Newark, NJ, 1994.
6. D. Tzovaras, M.G. Strintzis, H. Sahinoglou, ‘Evaluation of multiresolution block matching techniques for motion and disparity estimation’, Signal Processing: Image Communication, Vol. 6, 59-67, 1994.
7. R. Franich, R.L. Lagendijk, J. Biemond, ‘Stereo-enhanced displacement estimation by genetic block matching', SPIE Visual Communications and Image Processing, Vol.2094, pp.362-371, 1993.
8. M.G. Perkins, ‘Data compression of stereopairs’, IEEE Trans. on Communications, Vol.40, No.4, pp.684-696, April 1992.
9. S. Liu, M. Hayes, ‘Segmentation based coding of motion difference and motion field images for low bit-rate video compression’, Proc. of ICASSP ‘92, Vol.III, pp.525-528, 1992.
10. A. Tamtaoui, C. Labit, ‘Constrained disparity and motion estimators for 3DTV image sequence coding', Signal Processing: Image Communication, vol.4, pp.45-54,1991.
11. A. Tamtaoui, C. Labit, ‘Coherent disparity and motion compensation in 3DTV image sequence coding schemes', Proc. of ICASSP ‘91, Vol.IV, pp.2845-2848, 1991.
12. M. Waldowski, ‘A new segmentation algorithm for videophone applications based on stereo image pairs’, IEEE Trans. on Communications, Vol.39, No.12, December 1991.
13. I. Dinstein, et al., ‘Compression of stereo images and the evaluation of its effects on 3-D perception’, SPIE Applications of Digital Image Processing XII, Vol.1153, pp.522-529, 1989.
14. B. Julesz, ‘Foundations of cyclopean perception’, University of Chicago Press, 1971.