Residual image coding for stereo image compression

Tamás Frajka
Kenneth Zeger
University of California
Department of Electrical and Computer Engineering
San Diego, La Jolla, California 92093-0407
E-mail: frajka@code.ucsd.edu

Abstract. One main focus of research in stereo image coding has been disparity estimation, a technique used to reduce the coding rate by taking advantage of the redundancy in a stereo image pair. Significantly less effort has been put into the coding of the residual image. These images display characteristics that are different from that of natural images. We propose a new method for the coding of residual images that takes into account the properties of residual images. Particular attention is paid to the effects of occlusion and the correlation properties of residual images that result from block-based disparity estimation. The embedded, progressive nature of our coder enables one to stop decoding at any time. We demonstrate that it is possible to achieve good results with a computationally simple method. [DOI: 10.1117/1.1526492]

Subject terms: stereo image; residual image compression; progressive image coding.

Paper 020170 received May 1, 2002; revised manuscript received June 11, 2002; accepted for publication June 12, 2002.

1 Introduction

Human depth perception in part relies on the difference in the images the left and right eyes send to the brain. By presenting the appropriate image of a stereo pair to the left and right eyes, the viewer perceives scenes in three dimensions instead of as a 2-D image. Such binocular visual information is useful in many fields, such as telepresence style video conferencing, telemedicine, remote sensing, and computer vision.

These applications require the storage or transmission of the stereo pair. Since the images seen with the left and right eye differ only in small areas, techniques that try to exploit the dependency can yield better performance than independent coding of the image pair.

Most successful techniques rely on disparity compensation to achieve good performance. Disparity compensation is similar to motion compensation for video compression. It can be carried out in the spatial domain,[1–5] or in the transform domain.[6] Disparity compensation can be a computationally complex process. In Ref. 7, a wavelet-transform-based method is used for stereo image coding that does not rely on disparity compensation. While this technique is simpler, it sacrifices compression performance for reduced complexity.

Many of the preceding papers use discrete cosine transform (DCT)-based coding of the images, which uses a rate allocation method to divide the available bandwidth between the two images. For each target bit rate, a new optimization must be performed to find the optimal balance. Embedded coding techniques based on the wavelet transform[8,9] provide improved performance for still images when compared with DCT-based methods. An embedded bit stream can be truncated at any point to obtain the best reconstruction for the given bit rate. An embedded stereo image coding scheme is proposed in Ref. 10 that achieves good performance without having to use rate allocation.

With disparity compensation, one image is used as a reference image, and the other is predicted using the reference image. The gain over independent coding comes from compressing the residual image that is obtained as the difference of the original and predicted images. Little attention has been paid to the coding of the residual image. Moellenhoff and Maier[11] examined the properties of disparity-compensated residual images and proposed some DCT and wavelet techniques for their improved encoding.

In this paper, we propose a progressive coding technique for the compression of stereo images. The emphasis of our work is on the coding of the residual image. These images exhibit properties different from natural images. Our coding techniques make use of these differences. We show that the correlation across block boundaries in the residual image is diminished, suggesting that the coding of these blocks individually might be preferable. We propose to use transforms that take into account the correlation properties of the residual image as well as the block-based nature of the disparity estimator used in our coder. Occluded blocks are often difficult to estimate. Residuals of occluded blocks therefore are different from blocks that are well matched by the disparity estimation (DE) process. In our technique, we treat these two types of blocks differently. The image transform we propose uses a DCT on the blocks that are well matched by DE and Haar filtering on the occluded blocks, resulting in a mixed image transform. Multigrid embedding[12] (MGE) is used as the embedded image coder. It provides similar performance as zerotree-based techniques with increased flexibility. While the components of our proposed method are of low complexity, they yield some significant improvements over other methods in the coding of stereo images.

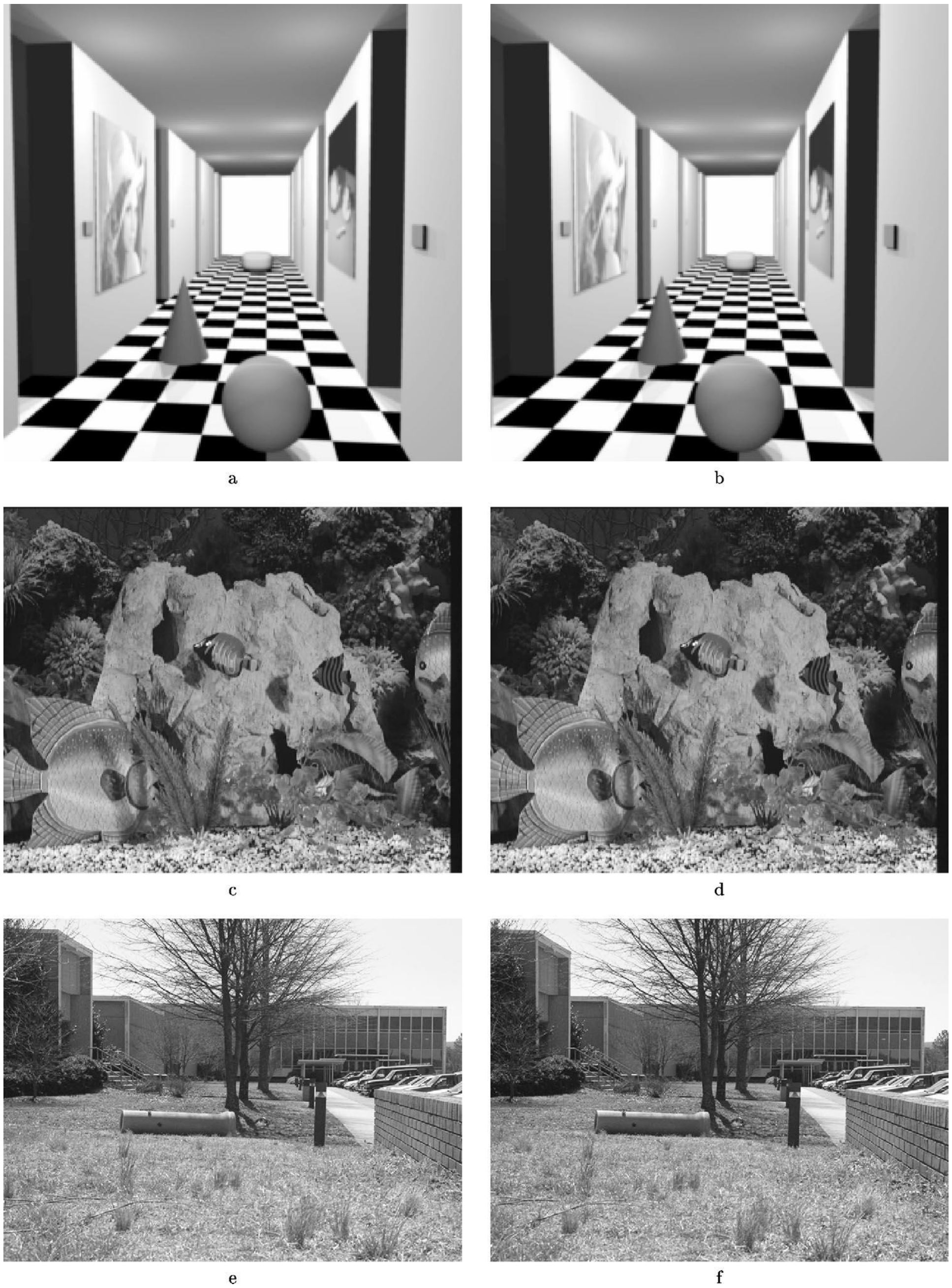


Fig. 1 Original images of the (a) and (b) ''Room,'' (c) and (d) ''Aqua,'' and (e) and (f) ''Outdoors'' stereo pairs.

The outline of the paper is as follows. Section 2 gives an overview of stereo image coding. Our contribution is in Sec. 3, with experimental results provided in Sec. 4. Finally, the conclusion is given in Sec. 5.

2 Stereo Image Coding

Stereoscopic image pairs represent a view of the same scene from two slightly different positions. When the images are presented to the respective eye, the human observer perceives the depth in the scene as in three dimensions. One can obtain stereo pairs by taking pictures with two cameras that are placed in parallel 2 to 3 in. apart. Figure 1 shows the original images of three different stereo pairs. The synthetic ''Room'' image represents such applications as video games or virtual reality, while the two natural scenes provide different distance scenarios, which exhibit different disparity properties. The left and right images of the ''Room'' pair differ mainly in the left edge of the left image where a piece of the wall is visible that cannot be found in the right image. Certain areas of the floor tile are differently covered by the cone and ball in these images. In the ''Aqua'' pair, the differences occur at the left and right edges of the images as well as around the rock in the middle. Because of the larger distance between

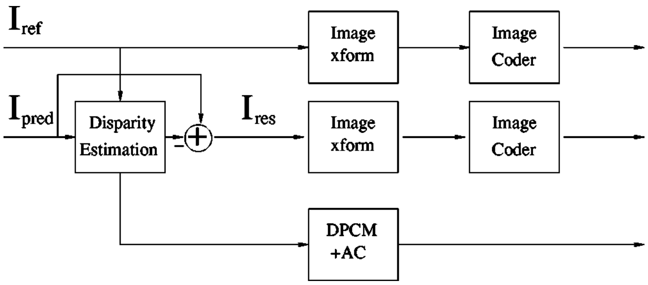


Fig. 2 Stereo image encoder based on disparity estimation.

the objects and the camera of the ''Outdoors'' pair, they exhibit the smallest differences, mostly around the image edges.

Because of the different perspective, the same point in the object will be mapped to different coordinates in the left and right images. Let (x_l, y_l) and (x_r, y_r) denote the coordinates of an object point in the left and right images, respectively. The disparity is the difference between these coordinates, $\mathbf{d}=(x_l-x_r, y_l-y_r)$. If the cameras are placed in parallel, then $y_l-y_r=0$, and the disparity is limited to the horizontal direction. One image of the pair serves as a reference image I_{ref} and the other I_{pred} is disparity estimated with respect to the reference image. A block diagram of a typical encoder using DE is shown in Fig. 2. This structure is often referred to as CONCOD (CONditional CODer). Perkins in Ref. 13 shows that in general this structure is suboptimal in the rate-distortion sense.

The disparity of each object in the image depends on the distance between the object and the camera lenses. Objects closer to the lens display larger disparity. As most natural images are composed of objects at different distances from the camera lens no single disparity value could be used to describe the difference between the left and right images (see Refs. 14 and 15 for more details.) The disparity estimation process must determine the correct displacement for each image pixel. Since this process would be quite complex if done for each pixel individually, it is usually carried out for groups of pixels instead. The simplest such grouping is the use of $k\times k$ nonoverlapping blocks. Block sizes using $k=8$ or $k=16$ provide a good trade-off between accuracy of the estimation and the number of bits necessary to encode the disparity vector $\mathbf{d}$ for each block.

The search for the matching block is carried out in a limited search window. Given the reference image, the optimal match could be any $k\times k$ block of the image. This exhaustive search is computationally complex. From the parallel camera axis assumption one can restrict the search to horizontal displacements only. (A small vertical displacement can also be allowed to compensate for the inaccuracy of practical camera systems.) From the camera setup it is clear that the disparity for objects in the left image with respect to the right image is positive and vice versa. This observation helps further limit the scope of the search.

Other techniques aimed at reducing the computational cost of full search DE include dynamic programming,[1] hierarchical DE (Ref. 4), and adaptive directional, limited-search algorithms.[2]

The estimation process works well for blocks that are present in both images. However, occlusion may result if certain image information is present only in one of the images. Occlusion can happen for two main reasons: finite viewing area and depth discontinuity. Finite viewing area occurs on the left side of the left image and the right side of the right image where each eye can see objects that the other eye cannot. Depth discontinuity is due to overlapping objects in the image; certain portions can be hidden from one eye on which the other eye has direct sight.

Another cause of mismatch is photometric variations. This phenomenon is due the variation of the reflected light that reaches the left and right lenses. A simple, global solution to this problem is histogram modification, as proposed in Ref. 16.

At the price of increased complexity, several methods were proposed that try to improve DE for both occlusion and photometric variations: subspace projection,[17] sequential orthogonal subspace updating,[18] and overlapped block DE (Ref. 19).

The disparity vectors are usually losslessly transmitted using differential pulse code modulation (DPCM) followed by arithmetic coding.[20] Tzovaras and Strintzis[21] proposed a rate-distortion framework for the encoding of the disparity vector field, allowing some distortion in the transmission of the displacement vectors.

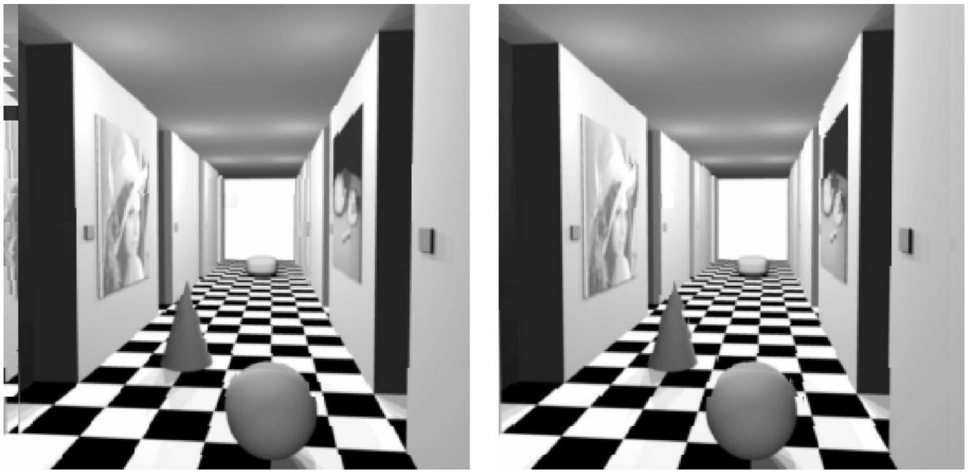
Given the disparity estimate of the image, the residual I_{res} is formed by subtracting the estimate from the original. This residual and the reference image are then encoded. Many proposed techniques use DCT-based block-coding methods for the encoding of both images. They also require a bit allocation mechanism to determine the coding rate of each image. (This bit allocation is carried out in addition to the bit allocation between the DCT-transformed blocks of each image.) For each target bit rate, a separate optimization is used to determine the appropriate bit allocation. Woo and Ortega[22] perform a blockwise-dependent optimization instead of independent optimization for the reference and residual images to improve the coding performance.

Embedded image coders can be terminated at any bit rate and still yield their best reconstruction to that rate without *a priori* optimization. Zerotree-style techniques such as the embedded zerotree wavelet (EZW) by Shapiro,[8] or set partitioning in hierarchical trees (SPIHT) by Said and Pearlman[9] offer excellent compression performance for still images. These zerotree techniques are extended to stereo images by Boulgouris and Strintzis.[10] The bit plane coding is performed on both the residual and reference image at the same time, guaranteeing that the most significant information for both images is sent before the less significant information.

The decoding of stereo images is straightforward. Both the residual and reference image are reconstructed. Using the DE information and the reconstructed reference image, the decoder can recover the other image of the stereo pair.

3 Residual Image Coding

The goal of this paper is to make stereo image coding more efficient by improving the coding of the residual image. The DE we chose is rather simple, but even with such a simple disparity estimator our proposed coding technique has very good performance.



a b

Fig. 3 (a) Estimate of left image with right image as reference and (b) estimate of the right image with left image as reference.

3.1 *Image Coding Method*

Embedded coding yields good performance coupled with simplicity of coding due to not having to perform any bit allocation procedure. MGE by Lan and Tewfik[12] uses a quadtree structure instead of the zerotrees of EZW or SPIHT. It uses the same bit plane coding, starting from the most significant bits of the transform domain image down to the least significant. For each bit plane, the quadtree structure is used to identify the significant coefficients, i.e., those whose most significant bit is found on that bit plane. The ''sorting'' pass identifies the coefficients that become significant on the current bit plane, while the ''refinement'' pass refines those coefficients that have previously become significant. Their results demonstrated that this technique outperforms the zerotree-based methods on images with significant high-frequency content. As residual images contain edges and other high-frequency information MGE is a natural candidate for their encoding.

The way we use MGE for stereo image compression is similar to that in Ref. 10. For each bit plane, first the sorting and refinement pass are executed for the reference image and then for the residual image. The highest magnitude coefficient is usually smaller for the residual image than for the reference image.

3.2 *Occlusion*

As noted in Sec. 2, there are two kinds of occlusion that may occur in DE. A finite viewing area can be overcome in certain cases. If a one-directional search is used (as suggested by the observation on the direction of the displacement in Sec. 2), that method could run out of image pixels at the edge of the image where it would also have difficulty finding the corresponding block in the reference image. If, however, we allow the search to continue in the other direction, it may find blocks similar to the one to be estimated. This can be seen in Fig. 3, where the estimate of the left image on the left edge clearly displays some occlusion error, while the right edge of the right image looks almost identical to the original.

The residual image of those blocks that are occluded because of depth discontinuity display different characteristics from the other parts of the image. As noted in Ref. 23, the occluded blocks are more correlated. We propose to detect such blocks, and code them differently from the rest of the residual image blocks for improved efficiency.

3.3 *Image Transform*

Moellenhoff and Maier's analysis[23] indicates that residual images show significantly different characteristics from natural images. Residual images mainly contain edges and other high-frequency information. The correlation between neighboring pixels is smaller as well. This suggests that transforms that work well for natural images may not be as effective for residual images.

In wavelet transform coding, one of the most widely used filters is the 9-7 filter by Antonioni et al.[24] It is preferred for its regularity and smoothing properties. With the image pixels less correlated in residual images shorter filters can better capture the local changes. For this reason we propose the use of Haar filters. These 2-tap filters take the average (low pass) and difference (high-pass) of two neighboring pixels. As our experimental results show, the use of Haar filters improves performance.

DE uses $k\times k$ size blocks to find the best estimates for the image. There is no reason to expect neighboring blocks to exhibit similar residual properties. For one block, the algorithm can find a relatively good match, while its neighbor could be harder to predict from the reference image.

Moellenhoff's results indicate that the pixels of the residual image are less correlated than those of the original image. But they do not reveal much about the local correlation of pixels, namely, across the $k\times k$ block boundaries. We investigate 1-pixel correlation on a more local scale in both horizontal and vertical directions. Instead of gathering these statistics for the whole image, we look only at the correlation between all pixels in the n'th column/row and its immediate neighbor in the $(n+1)$'th column or row of all $k\times k$ blocks for the case of horizontal or vertical corre-

Table 1 Comparison of 1-pixel horizontal correlation for pixels in a given column of an 8×8 block of the right image of the "Room" and "Aqua" stereo image pairs.

	1	2	3	4	5	6	7	8
"Room" original	0.93	0.94	0.96	0.96	0.97	0.95	0.94	0.94
"Room" residual	0.27	0.38	0.41	0.45	0.44	0.31	0.33	0.03
"Aqua" original	0.90	0.89	0.89	0.89	0.88	0.88	0.87	0.89
"Aqua" residual	0.24	0.25	0.22	0.23	0.25	0.23	0.26	0.12

lation, respectively. Note that the correlation between the k'th and $(k+1)$'th columns/rows gives the correlation just across the boundary between two neighboring blocks. Table 1 shows the 1-pixel correlation in the horizontal direction for the right image and its residual using blocks of size 8×8. (The trends are similar for vertical correlation and the left image as well.) It can be seen in Table 1 that the 1-pixel correlation drops significantly at the block boundary (column 8) in the residual image, supporting our assumption that different blocks exhibit different properties in the residual image.

Based on this observation we focus on block-based transforms that can better capture the differences between the blocks than a global transform, such as the wavelet transform that sweeps across the block boundaries. The DCT in practice is performed on $k\times k$ blocks. Its performance is diminished by the JPEG encoding method. However, if the DCT coefficients are regrouped into a wavelet decomposition style subband structure, as proposed in Ref. 25, and are encoded using an embedded coder, the performance approaches that of wavelet based methods. [This method is referred to as embedded zerotree DCT (EZDCT).]

None of the proposed image transforms so far take into account the effect of occlusion. For an occluded block, the best match can still be a very distorted one. In those cases, not using the estimate for the given block at all could be the best strategy. This is similar to coding an I block in the H.263 video compression algorithm. For each block, the estimator should decide if the best match is good enough. If not, the given block is left intact. This process creates a mixed residual image, with some parts having mostly edges and high-frequency information, and other parts blocks from the original image. For residual blocks that contain significant high-frequency information, a uniform band partitioning (such as with DCT) works better than octave-band signal decomposition (see Ref. 26), while octave-band decomposition is desirable in blocks of the original image.

Note that the Haar transform uses only two neighboring pixels to compute the low- and high-frequency coefficients, then moves on to the next pair. If the block size k is even, then starting at the left edge of the block, the Haar transform can be performed without having to include pixels from outside the block for the computation of Haar wavelet coefficients for all pixels in the block. Furthermore, this can be repeated up to $\lfloor \log_2 k \rfloor$ levels without affecting coefficients from outside the $k\times k$ block. Here we propose to use a mixed image transform. This transform consists of a Haar transform of three levels for occluded blocks and DCT for others with the DCT coefficients regrouped into the wavelet subbands to line up with the Haar-transformed coefficients.

4 Experimental Results

In our simulations we used the 256×256 "Room" stereo image pair, and the Y component of the color stereo image pairs "Aqua" (360×288) and "Outdoors" (640×480) shown in Fig. 1. The reference image was transformed using the 9-7 filters. For DE, a simple scheme was used with

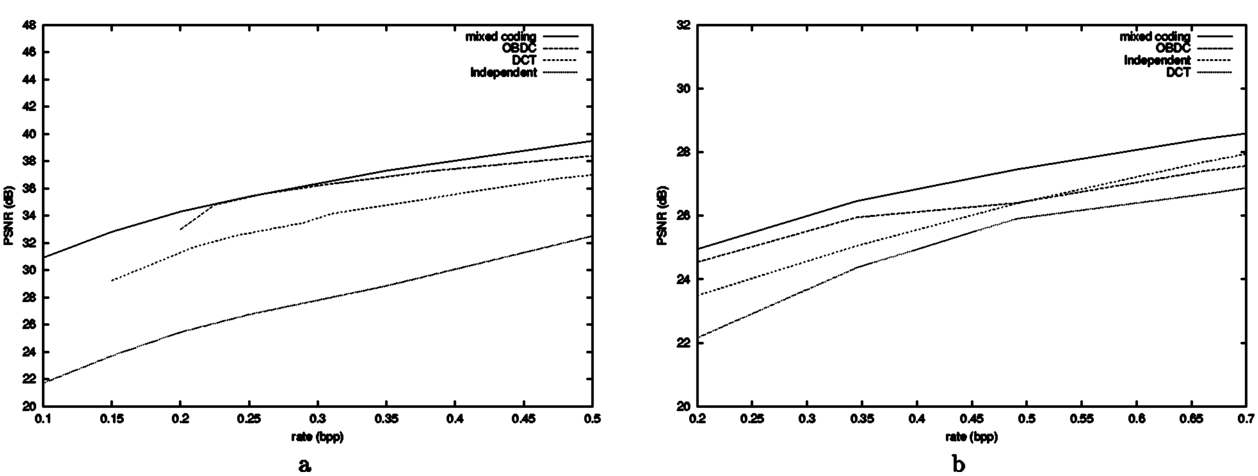


Fig. 4 Comparison of independent coding, JPEG-style coding, OBDC, and mixed transform coding for the left image residual (with reference image JPEG-coded with quality factor 75) for (a) the "Room," and (b) the "Aqua" images.

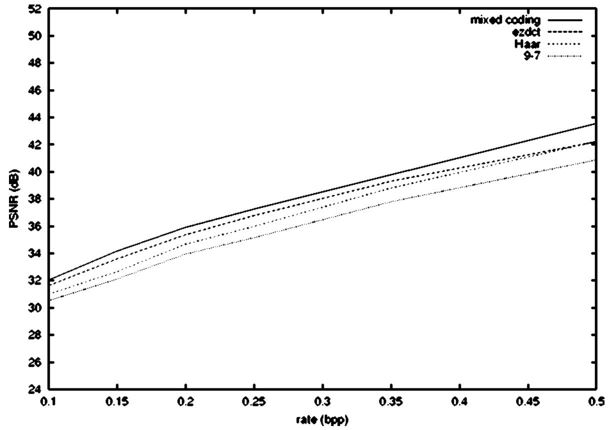


Fig. 5 Comparison of 9-7 wavelet transform, Haar transform, EZDCT-style, and mixed transform coding for the left image of the "Room" stereo pair.

a 64-pixel horizontal search window. Occlusion detection consisted of looking for blocks where the estimation error was above a given threshold.

We present our results both visually and in terms of peak signal-to-noise ratio (PSNR). For stereo images, the PSNR is computed using the average of the mean squared error (MSE) of the reconstructed left and right images,

$$\mathrm{PSNR} = 10 \log_{10} \frac{255^2}{(\mathrm{MSE}_l + \mathrm{MSE}_r)/2}.$$

First we compare different methods for the coding of the disparity estimated left image for the "Room" and "Aqua" pairs. The reference image is the JPEG coded (quality factor 75) right image. (This is chosen to be able to compare the results with previously published work.[19]) The bit rate figures include the coding of the disparity vector field. In the case of the mixed transform, for each block an extra bit is encoded using context-based arithmetic coding to signal if that block is considered as occluded. (In the case of independent coding, it is not necessary to encode any disparity information.) The PSNR is computed using the MSE for

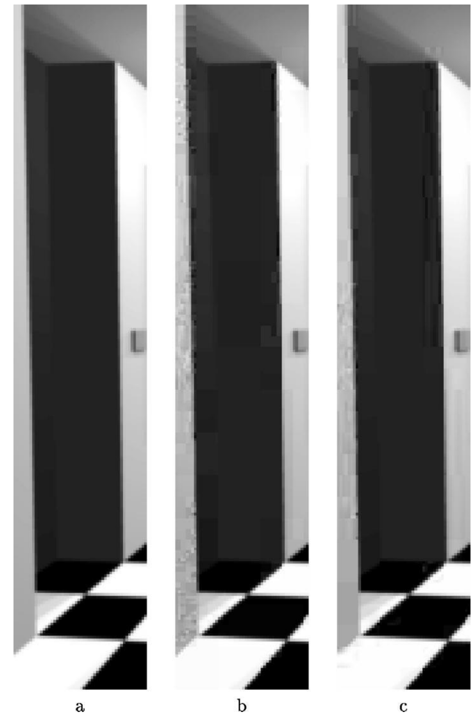


Fig. 7 Occluded area of left image; (a) original, (b) compressed with the Haar transform at 0.15 bpp, and (c) compressed with the mixed transform at 0.15 bpp.

the left image alone. The JPEG-style coder in our comparison uses quantization tables from the MPEG predicted frame coder.

Figure 4 compares independent coding, JPEG-style coding, overlapped block disparity compensation[19] (OBDC), and mixed transform coding. Mixed transform coding significantly outperforms both independent and JPEG-style coding with a gain of about 3 dB over the JPEG-style encoding. It also performs as well or better than OBDC coding, which uses a computationally more complex disparity estimator.

Figure 5 shows the effect of different coding techniques of the residual image and compares their performance. As

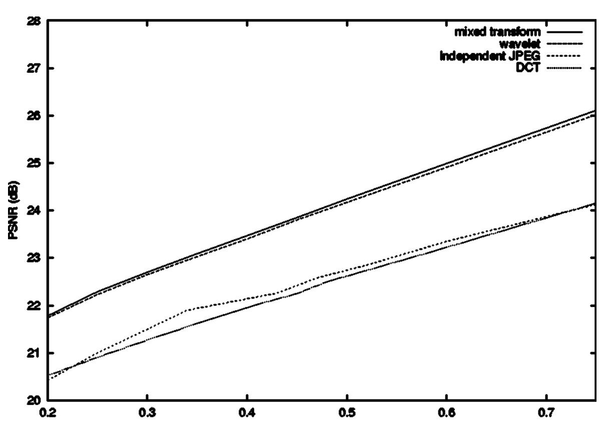


Fig. 6 Comparison of independent coding, JPEG-style image coding, wavelet transform, and mixed transform coding for the left image of the "Outdoors" stereo pair.

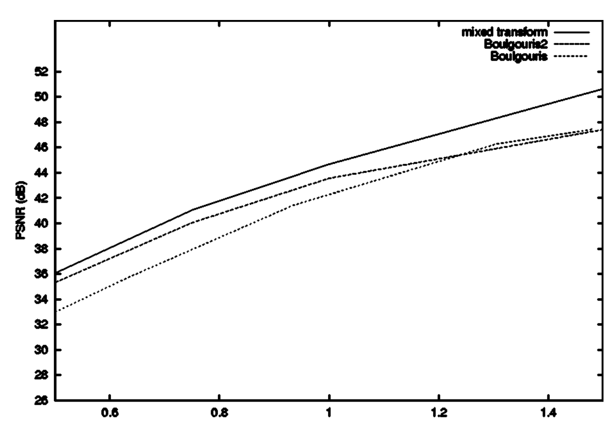


Fig. 8 Comparison of proposed method with the embedded stereo coding scheme from Boulgouris and its improved version for the full "Room" stereo pair.

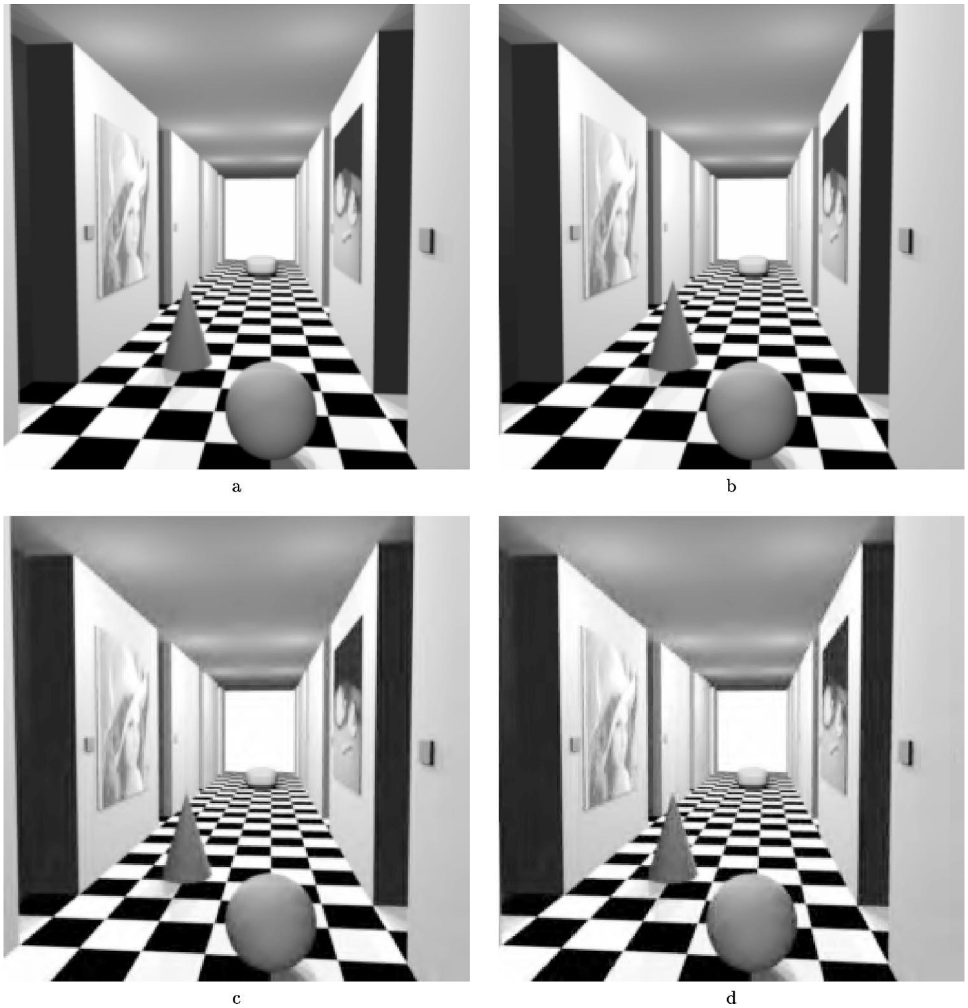


Fig. 9 (a) and (b) Original ''Room'' image pair and (c) and (d) mixed transform compressed version of the ''Room'' image at 0.5 bpp.

can be seen, the 9-7 wavelet filters perform poorest on the residual image, followed by Haar filtering, EZDCT-style coding, and mixed transform coding. The differences range from 0.5 to 1 dB for different bit rates between the pairs in the preceding ranking.

A similar comparison is given for the ''Outdoors'' image pair in Fig. 6. These images contain less occluded areas and the natural images are also harder to predict using such a simple disparity estimator. Using our mixed transform still produces up to 0.2 dB improvement over wavelet coding of the residual image. While the results for these two images improve the performance of DCT-based techniques they still fall short of the performance of individual wavelet coding of these images by about 0.2 dB.

Figure 7 demonstrates the effectiveness of the mixed transform. Occlusion in the left image occurs around its left edge. Figure 7(a) shows the original image, Fig. 7(b) shows the result compressed using the Haar-transform, and Fig. 7(c) the outcome after using the mixed transform. The wall area is more uniform in Fig. 7(c) because the mixed coder better preserved the occluded block.

For images that do not contain significant occluded information, the performance of the mixed transform coder is almost identical to that of the EZDCT-style coder.

Next we compare our proposed method and the results from Ref. 10 for the ''Room'' pair. Good residual image performance alone does not guarantee overall good performance when the entire stereo image is concerned in an embedded coding scenario. Recall that the decoder uses the compressed reference image to recreate the estimate for the other image. If the coding of the residual image takes away bits from the coding of the reference image the overall result may not be as good as the coding of the residual image would suggest.

Figure 8 demonstrates this comparison. In this case the left image is chosen as the reference image. In the comparison, ''Boulgouris2'' refers to new results (received from the authors of Ref. 10) obtained by an improved version of the original embedded stereo coder. It uses a more sophisticated disparity estimator and better wavelet filters. Our proposed method outperforms this improved algorithm as well by 0.70 to 2.3 dB.

The original ''Room'' stereo image pair and its mixed transform compressed version at 0.5 bpp is presented in Fig. 9. At this rate, there is little noticeable distortion between the original and the compressed images. To fully evaluate the effect of compression on stereo perception one would need a stereo viewer to fuse these images.

5 Conclusion

This work focused on the coding of the residual image in a stereo image compression scenario. Our method specifically addresses the issue of the image transform and the handling of the occluded blocks in the residual image. We showed that by individually coding the blocks of the residual image corresponding to the DE process we can take advantage of the correlation properties of residual images. Occlusion is handled by foregoing estimation for those blocks whose prediction is very distorted. Using an embedded encoding scheme enables the encoding to be stopped at any given rate without having to perform bit allocation. While the encoding is computationally simple, our simulations show improvements over previously published results.

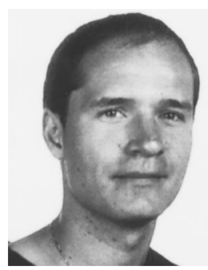
In future research, this work can be extended to investigating the properties of residual images that result from more sophisticated DE techniques and applying some of the proposed methods to improve their coding, especially for the case of natural images.

Acknowledgments

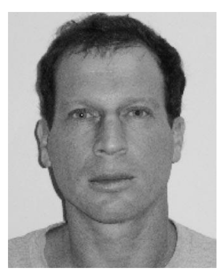
The authors would like to thank Eve Riskin for helpful comments and Nikolaos Boulgouris for providing data for the performance comparison. This work was supported in part by the National Science Foundation and the University of California, San Diego (UCSD), Center for Wireless Communications.

References

1. H.-H. Jeon, J.-H. Kim, and D.-G. Oh, ''Stereo image coding with disparity compensation using dynamic programming,'' in *Proc. Int. Symp. on Cons. Elec.*, pp. 214–217 (Dec. 1997).
2. W.-H. Kim, J.-Y. Ahn, and S.-W. Ra, ''An efficient disparity estimation algorithm for stereoscopic image compression,'' *IEEE Trans. Consum. Electron.* **43**(2), 165–172 (1997).
3. W.-H. Kim and S.-W. Ra, ''Fast disparity estimation using geometric properties and selective sample decimation for stereoscopic image coding,'' *IEEE Trans. Consum. Electron.* **45**(1), 203–209 (1999).
4. C.-W. Lin, E.-Y. Fei, and Y.-C. Chen, ''Hierarchical disparity estimation using spatial correlation,'' *IEEE Trans. Consum. Electron.* **44**(3), 630–637 (1998).
5. D. Tzovaras and M. G. Strintzis, ''Disparity estimation using rate-distortion theory for stereo image sequence coding,'' in *Proc. Int. Conf. on DSP*, Vol. 1, pp. 413–416 (July 1997).
6. Q. Jiang, J. J. Lee, and M. H. Hayes III, ''A wavelet based stereo image coding algorithm,'' in *Proc. ICASSP*, Vol. 6, pp. 3157–3160 (Mar. 1999).
7. W. D. Reynolds and R. V. Kenyon, ''The wavelet transform and the suppression theory of binocular vision for stereo image compression,'' in *Proc. ICIP*, Vol. 2, pp. 557–560 (Sep. 1996).
8. J. M. Shapiro, ''Embedded image coding using zerotrees of wavelet coefficients,'' *IEEE Trans. Signal Process.* **41**(12), 3445–3462 (1993).
9. A. Said and W. A. Pearlman, ''A new, fast, and efficient image codec based on set partitioning in hierarchical trees,'' *IEEE Trans. Circuits Syst. Video Technol.* **6**(3), 243–250 (1996).
10. N. V. Boulgouris and M. G. Strintzis, ''Embedded coding of stereo images,'' in *Proc. ICIP*, Vol. 3, pp. 640–643 (Sep. 2000).
11. M. S. Moellenhoff and M. W. Maier, ''Transform coding of stereo image residuals,'' *IEEE Trans. Image Process.* **7**(6), 804–812 (1998).
12. T. Lan and A. H. Tewfik, ''Multigrid embedding (MGE) image coding,'' in *Proc. ICIP*, Vol. 3, pp. 369–373 (Oct. 1999).
13. M. G. Perkins, ''Data compression of stereopairs,'' *IEEE Trans. Commun.* **40**(4), 684–696 (1992).
14. P. An, Z. Zhang, and L. Shi, ''Theory and experiment analysis of disparity for stereoscopic image pair,'' in *Proc. Int. Symp. on Intel. Multimedia, Vid. and Speech Proc.*, Vol. 1, pp. 68–71 (May 2001).
15. H. Yamaguchi, Y. Tatehira, K. Akiyama, and Y. Kobayashi, ''Stereoscopic images disparity for predictive coding,'' in *Proc. ICASSP*, Vol. 3, pp. 1976–1979 (May 1989).
16. I. Dinstein, G. Guy, J. Rabany, J. Tzelgov, and A. Henik, ''On stereo image coding,'' in *Proc. Int. Conf. on Pattern Recognition*, Vol. 1, pp. 357–359 (Nov. 1988).
17. H. Aydinoğlu and M. H. Hayes III, ''Stereo image coding: a projection approach,'' *IEEE Trans. Image Process.* **7**(4), 506–516 (1998).
18. S.-H. Seo and M. R. Azimi-Sadjadi, ''A 2-D filtering scheme for stereo image compression using sequential orthogonal subspace updating,'' *IEEE Trans. Circuits Syst. Video Technol.* **11**(1), 52–66 (2001).
19. W. Woo and A. Ortega, ''Overlapped block disparity compensation with adaptive windows for stereo image coding,'' *IEEE Trans. Circuits Syst. Video Technol.* **10**(2), 861–867 (2000).
20. I. H. Witten, R. M. Neal, and J. G. Cleary, ''Arithmetic coding for data compression,'' *Commun. ACM* **30**(6), 520–540 (1987).
21. D. Tzovaras and M. G. Strintzis, ''Motion and disparity field estimation using rate-distortion optimization,'' *IEEE Trans. Circuits Syst. Video Technol.* **8**(2), 171–180 (1998).
22. W. Woo and A. Ortega, ''Optimal blockwise dependent quantization for stereo image coding,'' *IEEE Trans. Circuit Syst. Video Technol.* **9**(6), 861–867 (1999).
23. M. S. Moellenhoff and M. W. Maier, ''Characteristics of disparity-compensated stereo image pair residuals,'' *Signal Process. Image Commun.* **14**(1–2), 55–69 (1998).
24. M. Antonioni, M. Barlaud, P. Mathieu, and I. Daubechies, ''Image coding using wavelet transform,'' *IEEE Trans. Image Process.* **1**(2), 205–220 (1992).
25. Z. Xiong, O. G. Guleryuz, and M. T. Orchard, ''A DCT-based embedded image coder,'' *IEEE Signal Process. Lett.* **3**(11), 289–290 (1996).
26. T. D. Tran and T. Q. Nguyen, ''A progressive transmission image coder using linear phase uniform filterbanks as block transforms,'' *IEEE Trans. Image Process.* **8**(11), 1493–1507 (1999).



Tamás Frajka received his MS degree in computer science from the Technical University of Budapest in 1995. He is currently finishing his PhD degree in the Department of Electrical and Computer Engineering at the University of California at San Diego, La Jolla, CA.



Kenneth Zeger received both the SB and SM degrees in electrical engineering and computer science from MIT in 1984, and both the MA degree in mathematics and the PhD in electrical and computer engineering at the University of California, Santa Barbara, in 1989 and 1990 respectively. He was an assistant professor of Electrical Engineering at the University of Hawaii from 1990 to 1992. He was in the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign, as an assistant professor from 1992 to 1995, and as an associate professor from 1995 to 1996. He has been in the Department of Electrical and Computer Engineering at the University of California at San Diego, as an associate professor from 1996 to 1998, and as a professor from 1998 to present. He received an NSF Presidential Young Investigator Award in 1991. He served as Associate Editor At-Large for the *IEEE Transactions on Information Theory* during 1995–1998, served as a member of the Board of Governors of the IEEE Information Theory Society during 1998–2000, and is an IEEE Fellow.