Pioneering Block Based Stereo Image CODEC in Wavelet Domain

E.A.Edirisinghe *, M.Y.Nayan, H.E.Bez
Department of Computer Science, Loughborough University, UK

ABSTRACT

In this paper we propose the wavelet domain implementation of our original pioneering block based stereo image compression algorithm and compare its performance with traditional, DCT based and state-of-the art, DWT based stereo image compression algorithms. Due to the special requirements of the pioneering block based CODEC and the properties of DWT based multi-resolution decomposition, the implementation of the original algorithm in the wavelet domain is not straightforward and thus provides knowledge and understanding of significant novelty. Experiments were performed on a set of eight stereo image pairs representing, natural, synthetic, in-door and out-door images. We show that for the same bit rates, objective quality gains of up to 5 dB (PSNR) are obtained as compared to the benchmark algorithms. One significant property of the proposed CODEC is its ability to produce reconstructed right images of up to 25dB at right image bit rates as low as 0.1 bpp. Significant gains in subjective image quality are also obtained as compared to benchmark methods.

Indexing terms: Stereo image compression, wavelet-blocks, predictive coding and disparity compensation.

1. INTRODUCTION

The recent advancement of auto-stereoscopic display technologies are currently driving stereo imaging into applications domains such as digital television, computer simulations and Internet technologies that have been so far dominated by the state-of-the-art monocular image and video coding techniques. Thus at present, efficient algorithms that are capable of coding and transmitting (storing) stereo images and image sequences are attracting significant research efforts from the stereo imaging research community. To this extent various coding algorithms based on block, object and region based predictive coding techniques have been reported in literature [3-15].

In parallel to the abovementioned shift of technology from monocular to binocular vision capability, the popular, base compression technology that was used in almost all image and video coding standards (baseline-JPEG, MPEGs and H26x), Discrete Cosine Transforms [1], are at present being steadily replaced by DWT based technologies that are capable of providing improved rate-distortion performance and additional functionality [2,13-20]. The culmination of research in this direction resulted in the recent standardization of JPEG-2000 [2] and wavelet domain implementations of MPEG-4 video coding standard. As most stereo image and sequence compression techniques that have been proposed in literature are based on DCT based technologies [3-12], currently there exists a trend in transfer of stereo imaging technology to DWT based techniques. To this effect few wavelet based stereo image compression algorithms have been reported in literature [13-15] during the past few years. Most of these algorithms are a straightforward application of the popular wavelet based monocular image compression algorithms to stereo image compression.

Following the abovementioned trends in technology and research directions, within the present research context, we propose a wavelet-based implementation for the PBDCPC algorithm, which was originally implemented in [5,6] using DCTs as the base technology. The PBDCPC algorithm is known for its ability to work under very low bit rate constraints [6], still producing predicted images of acceptable image quality. This is mainly due to the fact that it avoids the necessity of having to transmit the so-called, disparity vector field as overhead bits. We show that the change of base technology from DCT to DWT further improves its low bit rate performance. However, due to the specific requirements of the PBDCPC coding structure and properties of DWT based multi-resolution decomposition [13,19], the above transfer of base technology is not straightforward and thus involves additional design and development of considerable novelty. The original PBDCPC algorithm is summarized in section 2 for quick reference. Readers interested in its detailed design and behavior are advised to refer to the original work reported in [5,6].

Stereoscopic Displays and Virtual Reality Systems X, Andrew J. Woods, Mark T. Bolas,
John O. Merritt, Stephen A. Benton, Editors, Proceedings of SPIE-IS&T Electronic Imaging,
SPIE Vol. 5006 (2003) © 2003 SPIE-IS&T · 0277-786X/03/$15.00

For clarity of presentation, the rest of the paper is organized as follows. Apart from section 1, which is a general introduction to the context of the proposed research, section 2, summarizes the principles and the strict requirements behind the effective implementation of the PBDCPC algorithm. Section 3, discuses the DWT based multi-resolution decomposition of a stereo image pair and the formation of wavelet-block based structure, which forms the foundation for the proposed implementation. Section 4 discusses PBDCPC algorithm's implementation in wavelet domain by first discussing a Multi-Resolution Pioneering Block Search (MRPBS) procedure and subsequently by re-structuring the resulting Predictive Error (PE) image and the reference image within a wavelet-block structure [19,20] for entropy coding and transmission. In section 5, we compare the performance of the proposed CODEC with that of traditional DCT based DCPC schemes [3,4] and a more recently published DWT based DCPC scheme [20] and prove the novel algorithm's improved rate-distortion performance and improved range of performance. Finally in section 6, we conclude with suggestions for further improvements of the proposed algorithms and its possible use in stereo image sequence compression.

2. PIONEERING BLOCK BASED SEARCH (PBS)

Figure 1, illustrates the *pioneering block based search* (PBS) procedure that is used in PBDCPC algorithm [6].

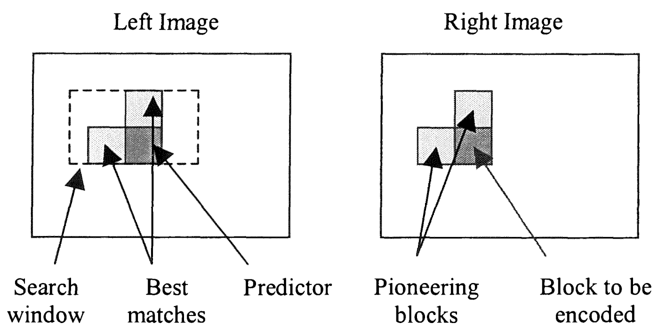


Figure 1. Pioneering block based search procedure

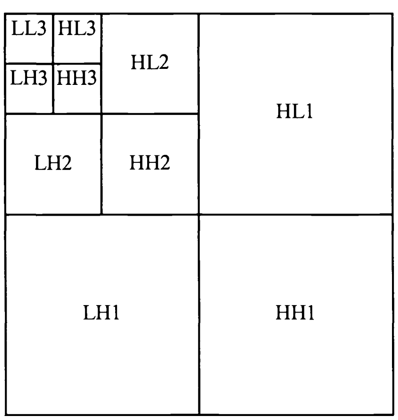
As illustrated in fig. 1, to encode a given block of image pixels (say size 8×8) in the predicted (say right) image, the block preceding it and the block directly above it are taken as *pioneering blocks* to search for a pair of blocks, within a selected windowed area of maximum likely of the locally reconstructed reference (left) image. The best matching pair is found using mean squared error as the minimization criterion and using equal weights for the two blocks. Once the best matching pair is found, the block directly below the upper block is chosen as the predictor for the block to be encoded. Note that at the boundaries of the image where one of the pioneering blocks (or both) is unavailable, the above search procedure is replaced by a PBS with one block (or by direct prediction). Once the predictor, $\hat{L}$ for the block to be encoded, R is found, the prediction error, E is calculated as, $E = R - \hat{L}$. Note the use of $\hat{L}$ to represent the fact that the search is performed on a locally reconstructed left frame. This is the first essential requirement of the PBDCPC coder, as it guarantees that the procedure adopted at the encoder end is identical to that adopted at the decoder end. This is due to the fact that at the decoder end what is available is the reconstructed reference frame. In addition to meeting the above condition, a second condition needs to be met. When the prediction error E is calculated at the encoder end, while it is being encoded and transmitted to the decoder end, the same coded error has to be decoded, and added to the predictor, $\hat{L}$, to produce the preceding pioneering block for the encoding of the next predicted frame block. Thus, always the pioneering block selection in the predicted image is done on a 'so-far' reconstructed image. This satisfies the second condition in guaranteeing that the encoder and decoder processes are identical. Thus in the PBDCPC coder the encoder and decoder work on identical reference frames and selects pioneering blocks from identical, 'so-far' reconstructed predicted frames. This enables the guaranteed reconstruction of the predicted image at the decoder end, without recourse to disparity information. We refer readers interested in the theoretical proof of the above-mentioned quality guaranteeing procedure to [21].

Following the pixel domain PBS procedure, summarized above, the actual prediction (i.e. calculation of prediction error, *E*) could be either performed in the pixel domain itself, or in the transformed domain. However, in order to satisfy the second condition in guaranteeing the reconstruction of the predicted image, the prediction errors have to go through transform coding/decoding and quantization/de-quantization at the encoder end, thus making it an integral part of the PBDCPC algorithm. Therefore in any attempt to change the base transform technology of PBDCPC algorithm to DWT the above-mentioned requirement takes precedence. Thus, a straightforward block based, pixel domain disparity compensated prediction to obtain the prediction error image in complete, and the subsequent DWT based coding of the same, is not a suitable DWT implementation of the PBDCPC algorithm. This is due to the fact that with such a scheme one is unable to introduce the compression losses that would be introduced to the prediction errors, before they are 'fed-back' to form the pioneering blocks for the next block to be encoded. Thus within our present research context we propose a strategy based on the formation of a wavelet-block coding structure, which removes the major obstacle faced by the abovementioned straightforward strategy.

3. WAVELET DECOMPOSITION OF A STEREO IMAGE PAIR & THE FORMATION OF WAVELET BLOCKS

The first stage of the proposed DWT implementation of the PBDCPC algorithm is the independent wavelet decomposition of the stereo image pair into N levels, forming 3N + 1 sub-bands. The coefficients of these sub-bands are generated by applying a cascade of two-channel filter banks to the image. Figure 2 illustrates this decomposition structure for a given image, with N=3. We name the lowest resolution band, LL3.

As described above the common practice in DWT subband coding procedures, is to group the coefficients into subband-oriented groups. In other words, coefficients of a given subband, comes from different spatial locations of the image. However in block based, DCT subband coding procedures, the coefficients are grouped into blocks, i.e. they are from common spatial locations, but different subbands. Since the adaptation of a DCT-like (baseline-JPEG like) coding structure would enable the proper design of the PBDCPC algorithm in the wavelet domain, we propose the transformation of the decomposition structure of figure 2 to wavelet-blocks as depicted in figure 3. The idea behind wavelet blocks is to group the DWT coefficients into blocks, so that the grouping is similar to that used by a DCT based subband coding procedure.



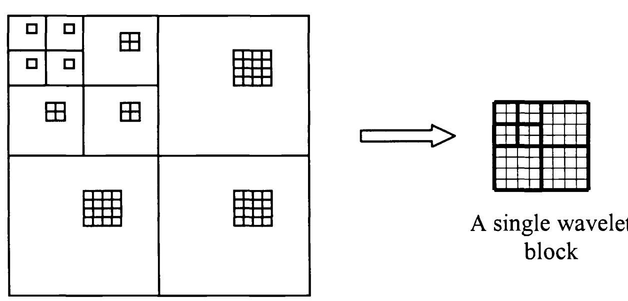


Figure 3. Formation of a single wavelet block from a 3-level decomposed image

Figure 2. Three level wavelet decomposition of an image

As illustrated in fig.3, the final result is the transformation of the input image into a re-organized structure of original image size, but consisting of a two-dimensional array of non-overlapping wavelet blocks, each representing different subband elements from the location corresponding to the block. In other words, for a *S*-level DWT, blocks of $2^S \times 2^S$ samples are constructed.

4. WAVELET BASED IMPLEMENTATION OF THE PBDCPC ALGORITHM

Figure 4 illustrates a high-level block diagram of the proposed CODEC. The original left image (reference) L_{ori}, is independently coded using a JPEG-like wavelet CODEC proposed in [19]. After reconstruction at the decoder end, it provides a reference for the prediction of the right image. In addition to the above, the encoded reference image is locally decoded at the encoder end to produce a reconstructed left image L_{rec}, which is subsequently used as the reference in the wavelet based PBS unit. Note that by adapting this strategy we satisfy the first condition of guaranteeing the reconstruction of the predicted image, under the pioneering block based coding technique. The original right image R_{ori} and reconstructed right image L_{rec} then undergo *N* level dyadic wavelet decomposition as described in section 3. For our experiments we have used *N*=3 and we follow the convention depicted in figure 2, in naming the ten resultant subbands, with the lowest subband being named as *LL3*.

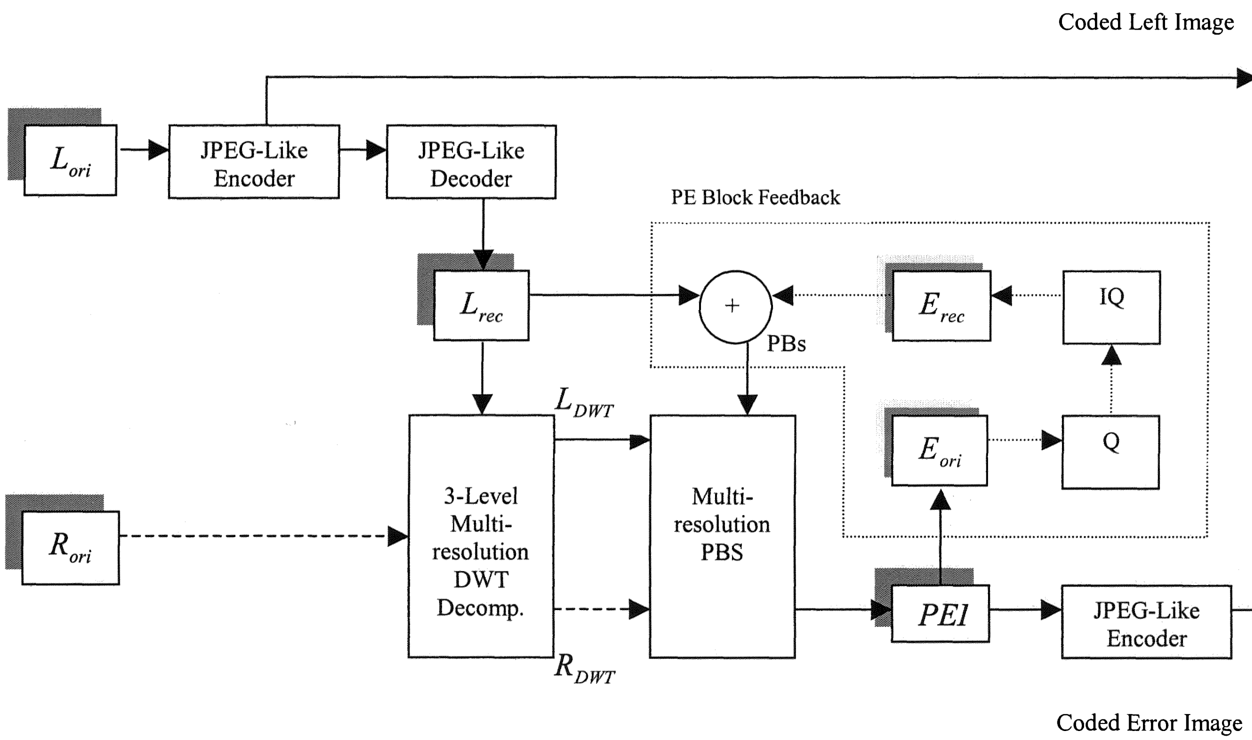


Figure 4. Proposed Encoder

Within the wavelet based PBS unit, for each corresponding subband pair (e.g. L_{LL3} and R_{LL3}, L_{HL1} and R_{HL1}) a PBS is performed as described in section 2, to produce the corresponding subband of the Prediction Error Image (PEI). Once the prediction error image has been completely found (i.e. all of its subbands found), a JPEG-like CODEC (see figure 5, for details refer [19]) is used to transform them into wavelet blocks and transmit after suitable quantization and entropy coding. Fig. 4 also illustrates a prediction error (PE) block feedback loop, which converts the original domain prediction errors into the reconstructed domain by locally sending the PE through a quantization/de-quantization procedure. The error block, in its reconstructed state, E_{rec} is then added to the best predictor previously found from L_{rec} to form the block, which would act as the preceding (front) pioneering block for the next iteration of the PBS. Thus the purpose of the above mentioned error feedback loop is to satisfy the second requirement for a successful PBDCPC CODEC design (see section 2).

The decoder of the proposed CODEC works in a manner similar to the encoder. Due to the orientation of the two pioneering blocks relative to the block to be encoded, in encoding blocks in the first row of blocks (other than the first block itself), of a given subband, only the block above is used as the pioneering block. A similar strategy in adapted to resolve the problem of encoding blocks in the first column of blocks. The block at the top, left-hand corner of a sub-band is always coded directly, i.e. without any disparity compensation.

5. EXPERIMENTAL RESULTS & ANALYSIS

The proposed CODEC was implemented using purpose built MATLAB routines and was tested on a set of eight commonly used test stereo image pairs representing indoor, outdoor, natural and synthetic scenes. All test stereo image pairs were of size 512×512. As parallel axis camera geometry has been used to acquire these images we have limited the PBS along a horizontal direction only. For both CODECs, a windowed search area representing a disparity range of 0-7 pixel shifts was used. For the benchmark CODEC, the disparity vector field was coded using a fixed length code, requiring three bits per block. A Daubachies compactly supported orthonormal filter (db7) was used as the wavelet transform.

In order to evaluate the performance of the proposed CODEC, in figure 5, we compare its performance against that of the well-known benchmark algorithm of Perkin's (benchmark-1), that uses DCT as the basic compression engine, and that of our previously proposed CODEC of [20] (benchmark-3), which is essentially a wavelet-block implementation of benchmark-1.

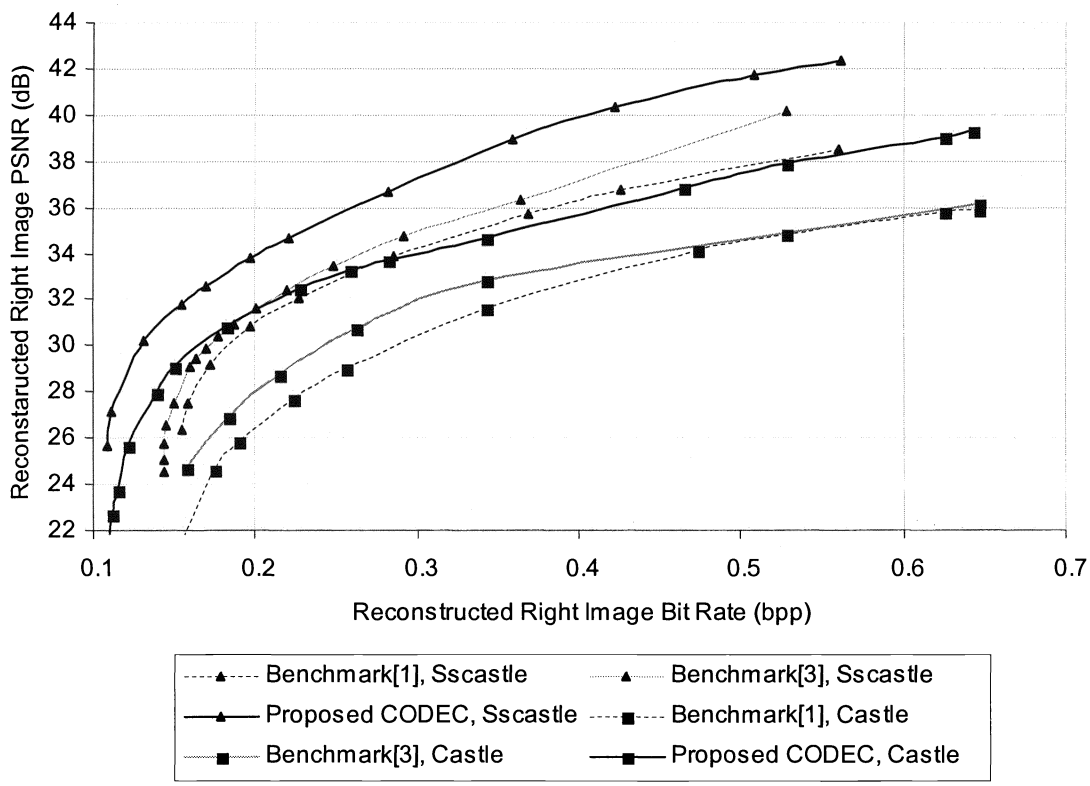


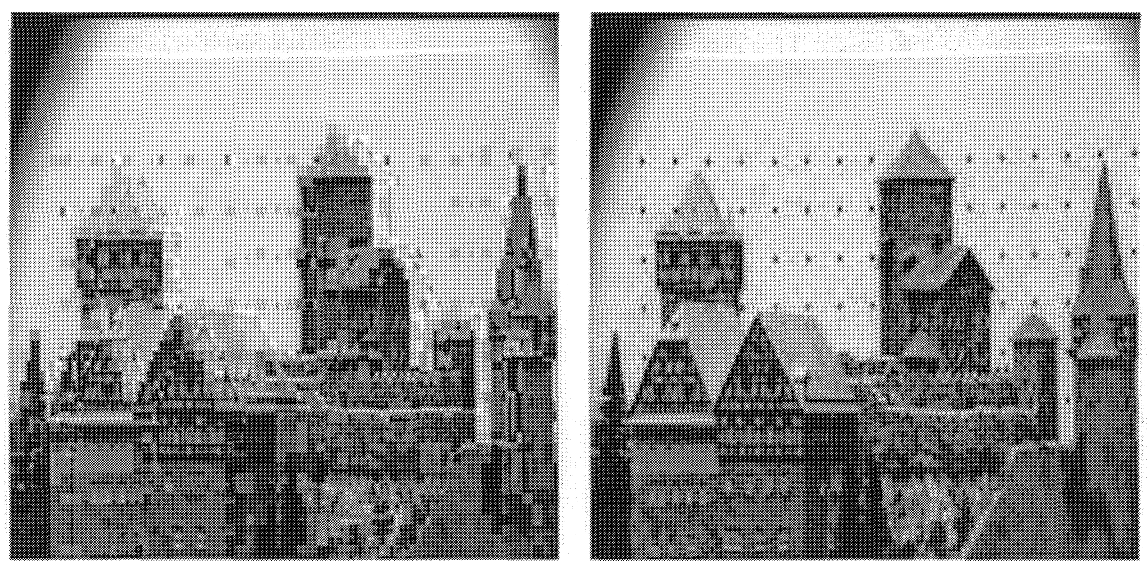
Figure 5. Comparison of rate-distortion performance graphs

The objective image quality is measured in terms of Peak Signal to Noise Ratio (PSNR), whereas the compression efficiency is measured in bits per pixels (bpp). The compression efficiency (or Bit Rate) is measured as follows:

$$BR = \left[\frac{Tot\_Bits_{comp}}{N \times M} \right] \quad bpp$$

where, N and M represent the image dimensions and Tot_Bits_{comp} is the total number of bits required to represent a given image in its compressed format. For benchmark-1 and benchmark-3, Tot_Bits_{comp} include the bits that are required to code the disparity vector field, in addition to the bit requirement to encode and transmit the predictive error image. However, for the proposed CODEC, Tot_Bits_{comp} represents only the bit requirement to encode the predictive error image. For a fair comparison, for a given stereo image pair, the objective quality of the left image (reference) was set at equivalent levels for all three CODECs.

As illustrated in fig. 5, the rate-distortion graphs show that the proposed PBDCPC CODEC outperforms the benchmark algorithms by considerable margins both at medium and low bit rates. At a reconstructed right image bit rate of 0.2 bpp, for Castle image, the PBDCPC CODEC produces a PSNR gain of approximately 3.5 dB as against benchmark-3 and 5 dB as against bechmark-1. For Sscastle image improvements of up to 3 dB are indicated at this bit rate. At higher bit rates, improvements in the range of 2-4 dB are indicated. Further analysis of the results in fig. 5 indicates the ability of the proposed PBDCBC CODEC to perform at bit rates as low as 0.11 bpp. In contrast, the benchmark CODECs are unable to perform at bit rates lower than 0.145 bpp. This is due to the fact that both benchmark CODECs require an allocation of a certain fixed bit budget (0.047 bpp in our experiments) for coding the resultant disparity vector fields.



(a) (b)

Figure 6. Subjective quality evaluation for Castle reconstructed right image (a) Benchmark-1, 0.1550 bpp, 21.29 dB (b) Proposed CODEC, 0.1498 bpp, 29.03 dB

Figure 6 compares the subjective image quality performance of the proposed CODEC against the DCT based Benchmark-1 at a bit rate of 0.15 bpp. It clearly illustrates that in the proposed CODEC the image quality near object boundaries are much better as compared to that of Benchmark-1.

In predictive coding of stereo image pairs, the predictive errors are highest near object boundaries, which represent areas of binocular disparity. Therefore the prediction error frames have highest values in these regions. Thus when these frames are coded under constrained bit budgets, the maximum loss is introduced in object boundary areas. This is particularly true at bit rates lower than 0.25 bpp for most images. Even though the proposed CODEC does not result in a more accurate prediction, the bit budget that could be allocated to coding the prediction errors is higher as compared to benchmark techniques 1 and 3 due to the fact that no bits have to be allocated for the coding of disparity vector fields. Particularly at very low bit rates, the bits required for the lossless coding of disparity vector fields is significantly higher than what is required for coding the prediction error fields. Thus the proposed CODEC's relative performance as compared to the benchmark techniques improves with the decrease of bit rate. A closer analysis of the fig. 6(a) indicates that the subjective quality degradation near object boundaries is of a *blocky* nature and in fig 6(b) it is of *ringing* nature. This is due to the inherent characteristics of DCT and DWT based compression. However, due to the usage of the pioneering block technique, in figure 6(b) the otherwise expected significant ringing effects have been largely eliminated.

6. CONCLUSION & FUTURE

In this paper we have proposed an efficient DWT implementation of our previously proposed DCT based pioneering block based stereo image compression algorithm. We have shown that due to the specific requirements of the design of the CODEC to guarantee the reconstruction of predicted images, without the use of disparity information, the DWT implementation is not a straightforward conversion of the base technology from DCT to DWT. The novel DWT based PBDCPC scheme results in improved, objective rate-distortion performance and a subjective quality performance. PSNR improvements of up to 5.5 dB are reported for transmission bit rates of 0.15 bpp. In addition to this the proposed CODEC is able to operate at bit rates as low as 0.1 bpp, where other CODECs fail to operate due to the need of a certain minimum bid budget for encoding disparity vector fields.

Further improvements to the proposed technique, in terms of algorithmic and computational efficiency is possible by using faster multi-resolution search techniques [13] that avoid searches in all sub-bands but would rather effectively re-use search information at lower resolution sub-bands within the search operations of higher resolution sub-bands. We are currently in the process of expanding the ideas proposed in this paper to wavelet based stereoscopic sequence coding.

REFERENCES

1. W.B.Pennebaker, J.L.Mitchell, "JPEG: Still Image Compression Standard", Van Nostrand Reinhold, NY, 1993.
2. ISO/IEC JTC 1/SC 29/WG 1, ISO/IEC FCD 15444-1: Information technology – JPEG-2000 Image Coding System: Core coding system, March 2000, www.jpeg.org/FCD15444-1.htm
3. M.G.Perkins, "Data Compression of Stereopairs", IEE Trans. on Comm., Vol.**40**, No.**4**, pp. 684-696, April 1992.
4. I.Dinstein, M.G.Kim, A.Hanik, "Compression of Stereo Images Using Subsampling and Transform Coding", Journal of Optical Engineering, Vol.**30**, pp. 1359-1364, Sept. 1991.
5. J.Jiang, E.A.Edirisinghe, H.Schroder, "A Novel Predictive Coding Algorithm for 3-D Image Compression", IEEE Trans. on Cons. Elec., Vol.**43**, No.**3**, pp. 430-437, August 1997.
6. J.Jiang, E.A.Edirisinghe, H.Schroder, "Algorithm for the Compression of Stereo Image Pairs", IEE Electronic Letters, Vol.**33**, No.**12**, pp. 1034-1035, June 1997.
7. W.Woo, A.Ortega, "Optimal Blockwise Dependent Quantization for Stereo Image Coding", IEEE Trans. on CSVT, Vol. **9**, No.**6**, pp. 861-867, Sep. 1999.
8. W.Woo, A.Ortega, "Overlapped Block Disparity Compensation with Adaptive Windows for Stereo Image Coding, "IEEE Trans. on CSVT, Vol. **10**, No.**2**, pp.194-200, March 2000.

9. N.D.Doulamis, A.D.Doulamis, Y.S.Avrithis, K.S.Ntalianis, S.D.Kollias, “Efficient Summarization of Stereoscopic Video Sequences”, IEEE Trans. on CSVT, Vol. **10**, No.**4**, pp. 501-517, June 2000.
10. R.Wang, Y.Wang, “Multiview Video Sewquence Analysis, Compression and Virtual Viewpoinr Synthesis”, IEEE Trans. on CSVT, Vol. **10**, No.**3**, pp. 397-410, April 2000.
11. N.Grammalidis, D.Beletsiotis, M.G.Strintzis, “Sprite Generation and Coding in Multiview Image Sequences”, IEEE Trans. on CSVT, Vol. **10**, No.**2**, pp. 302-311, March 2000.
12. J.Jiang, E.A.Edirisinghe "A Hybrid Scheme for Low Bit-Rate Coding of Stereo Images", IEEE Transactions on Image Processing, Vol. **11**, Issue. **2**, pp. 123-134, Feb. 2002.
13. S.Sethuramam, M.W.Siegel, A.G.Jordan, “A multi-resolution framework for stereoscopic image sequence compression’, Proceedings of the ICIP-95, Vol.2, pp. 361-365, 1995.
14. N.V.Boulgouris, M.G.Strintzis, “Embedded Coding of Stereo Images”, Proc. of the IEEE Int. Conf. on Image Proc., ICIP 2000, June 2000.
15. E.A.Edirisinghe, “Rate Scalable Stereo Image Coding Using EZW Algorithm”, Proc. of the IEEE Int. Conf. on Information, Communications and Signal Proc., ICICS 2001, Singapore, Oct. 2001.
16. S.A.Martucci, I.Sodagar, T.Chiang, Y.Zhang, “A Zerotree Wavelet Video Coder”, IEEE Trans. on CSVT, Vol.**7**, No.**1**, pp. 109-1997, February 1997.
17. K.Shen, E.J.Delp, “Wavelet Based Rate Scalable Video Compression, “IEEE Trans on CSVT, Vol.**9**, No. **1**, pp. 109-122, Feb. 1999.
18. S.A.Martucci, I.Sodagar, “Zerotree Entropy Coding of Wavelet Coefficients for Very Low Bit Rate Video Coding”, Proc. of 1996 IEEE Int. Conf. on Image Proc., Lausanne, Switzerland, Sept. 1996.
19. R. de Queiroz, C.K. Choi, Y. Huh, K.R. Rao, “Wavelet Transforms in a JPEG – Like Image Coder,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. **7**, No. **2**, pp 419 -424, April 1997.
20. M.Y.Nayan,E.A.Edirisinghe,H.E.Bez, “Baseline JPEG-Like DWT CODEC for Disparity Compensated Residual Coding of Stereo Images”, Proc. of Eurographic 2002, UK, pp. 67-74, June 2002.
21. Edirisinghe, E.A., Jiang, J., "Towards Eliminating Disparity Field Coding of Stereo Image Pairs", Proc. of the IEEE Int. Conf. on Information, Communication & Signal Processing, December 1999.

---

* Correspondence: E.A.Edirisinghe@lboro.ac.uk; Telephone: +44 (0)1509 228234; Fax: +44 (0)1509 211586