Vision, Video and Graphics (2005)
E. Trucco, M. Chantler (Editors)

Multi-View Image Coding with Wavelet Lifting Scheme

N. Anantrasirichai, C. Nishan Canagarajah, and David R. Bull

Department of Electrical & Electronic Engineering, University of Bristol, Bristol, UK

Abstract

In this paper, we present a novel multi-view image codec based on a wavelet lifting scheme. The proposed algorithm applies the disparity estimation and compensation with the lifting scheme. It is very efficient in terms of compressions performance, memory requirements and fast computation. To get the highest multi-view image coding efficiencies, two hybrid predictions are proposed; one for effective compression performance and one for compromising when more exact disparity values are required. Moreover, an adaptive weighing in update step and overlapped block disparity compensation are included to yield significant improvements in rate distortion performance. Experimental results show image quality gains of up to 2 dB and 1 dB against using well established methods such as the block-matching Haar and 5/3 wavelet lifting respectively.

1. Introduction

In the growing era of telecommunication, the high ability of storage and real-time transmission in image/video sequences has been dramatically grown in several manners so as to support a multiplicity of useful applications. Moreover, nowadays, various multimedia commercial systems are available for enthusiastic users to attain their demands on realistic visual communication. A number of three-dimensional (3D) visual technologies have been tried to meet these requirements, such as holography, two-view stereoscopic systems with special glasses and 3D wide screen cinema. Evidently, the multi-view image/video signals can provide the benefits of added realism, selective viewing, and improved scene comprehension. However, as the results of having increased the numbers of cameras, a three-dimensional imaging system results in a manifold spread in bandwidth over the existing monoscopic channel bandwidth. Consequently, this system has limited many commercial applications, such as internet video, digital television broadcasting and digital camcorders. Thus, the efficient compression algorithms are vital to reduce the size of data without sacrificing the perceived quality.

Fortunately, on account of the great correlation of stereoscopic views, diverse techniques have been proposed to exploit this characteristic. The first developed procedure is to code sum and difference of the two images in a stereo pair [Per92]. Afterward 3D discrete cosine transform coding of stereo image was presented [DGR*88]. It is equivalent to the sum-difference coding in transform domain. The performance of these two techniques decreases with increased disparity values by assuming that objects in the scene have same disparity costs. Nonetheless, these methods are not principally efficient since objects in the scene have normally different disparity values. Hence, the concept of disparity compensation, which established correspondence between similar areas in a stereo pair using binocular disparity information, was constituted [Luk96]. This method is used to predict the rest of the views from an independently coded view. In 1995, the results of MPEG-2 compatible coding have been presented. In the standard, one view is coded as the base layer and another view is coded within enhancement layer of the temporal scalability model and exploits discrete cosine transform (DCT) coding.

Recently, the wavelet scheme has been proposed for image/video coding. This approach is very efficient in terms of implementation and scalability properties. In this work, we extend this approach for multi-view image compression to remove spatial, temporal and geometric redundancies presenting in such sequences.

Moreover, the performance of the wavelet scheme can be enhanced by using an advanced prediction scheme to minimize correlation between views/frames. In temporal axis, the motion compensated temporal filtering (MCTF) was proposed by Ohm [Ohm94]. The block-based motion compensation is applied to three-dimensional discrete wavelet transform (3-D-DWT) and showed that a non invertible transform will affect coding gain if there are numerous disconnected pixels between blocks. This could be a significant problem in block-matching in disparity field due to some occlusion. Accordingly, the wavelet lifting scheme has been attractive, due to the invertible property. In [PB01], a non-linear lifting framework for temporal wavelet transform has been introduced which is shown to be superior to the MCTF. Luo et al. have implemented a motion compensated lifting structure by using bi-directional motion estimation in each lifting step of temporal direction [LLL*01]. Secker and Taubman have developed the highly scalable video coding approach [ST03]. It is based on the lifting with adaptive motion compensation. However, they have ignored the possibility of occlusion which is most difficulty in multi-view image system.

The implementation of the wavelet lifting scheme for multi-view image compression is novel and has not been investigated previously. In this paper, we present a framework which exploits spatial correlation as well as high geometric correlation present among multi-view images based on the wavelet lifting scheme. The paper is organized as follows: Section 2 briefly discuss about the importance of disparity estimation for multi-view system, whist Section 3 explains the fundamentals of wavelet lifting scheme and disparity compensation. The proposed codec is described in Section 4. The experimental results are presented in Section 5 with conclusions and future works in Section 6.

2. Disparity Estimation in Multi-View system

In the visual reality, at least two images taken in a bit different locations are required to perceive a 3D scene. This system contains a large amount of data comparing to the conventional monoscopic system that, unfortunately, is not suitable to store or transmit. Opportunely, these sequences have significant correlation among views, especially between the successive views, because of having been taken from the same scene. Hence, one view can be predicted by exploiting the corresponding information of other views and only the difference and side information are coded and transmitted. The task of finding the pairs of matching pixels is known as the correspondence problem and the distance between the two corresponding points when the two images are aligned one on top of the other is called the disparity.

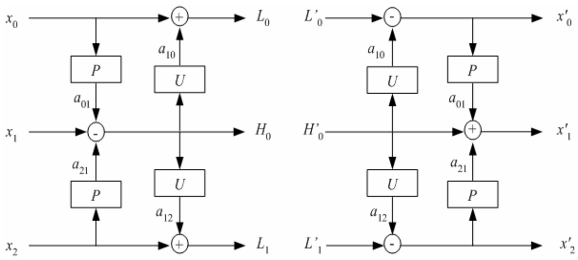
In 3D rendering, the virtual views make the smooth motion parallax feasible, so it should be similar to the real scene as much as possible. To synthesize the virtual view, at least two images, basically left image and right image, are required. However, it is problematic due to occlusion leading to the wrong depth information. That is, stereoscopic images make a viewer discomfort because the viewer prefers to vary degree of depth perception and the range of depth presenting in the scene. On the contrary, a greater sense of depth is provided by a relatively large inter-camera separation. The multi-view system offers images from several angles; hence the occlusion could be compensated by other views so that the correct depth and excellent synthesizing view are achieved [MSJ96]. That means the multi-view system supports look-around capability [ZFH*98]. Obviously, when the occlusion regions are correctly detected, the estimated disparities are more reliable because the erroneous disparity values of the regions that cannot be matched to any area in view references will not be generated. This exactly gives the advantage for precise calculation of the displacement between the intermediate view and other existing views in the view synthesis method.

The correspondence problem or the disparity estimation problem has been studied extensively. There are two main schemes of disparity estimation; correlation-based approach and feature-based approach. In the first one, the correlation-based approach searches its corresponding point within a search region and the disparity is the displacement when the correlation gets maximum cost. In other words, the disparity is found by minimizing the sum of error from predicting scheme. This method is simple but it could be highly efficient if some principles are satisfied, such as the light balance between views, the very small amount of dissimilarity or distortion between the views and etc. In feature-based approach, the disparity is searched by measuring the similarity of objects. It is more complicated than the first one, but it is more robust and provides algorithmic flexibility to the programmers [LT98].

However, the best algorithm for disparity estimation in multi-view image coding is trade-off between the limitation of required data size and the perfect virtual view construction. Some disparity estimation algorithms support the compression requirement but not suitable for being exploited in view synthesis. Some enhance approaches offer the correct depth results but it produces the massive data that affects the compression efficiency. Therefore, the appropriate method is selected upon the main purpose of the application.

3. Disparity Compensated Channel Filtering

The lifting scheme, one of various techniques exploited to construct wavelet bases or to factor existing wavelet filters into basic building blocks, was firstly introduced by Sweldens [Swe98]. Basically, the forward wavelet lifting method decomposes wavelet transforms into a set of stages. The operation starts with a split step, which divides the data set, *x*, into two groups. Normally, odd element, x_{2k+1}, and even element, x_{2k}, are chosen, because they are supposed to be similar and contain largest correlation. The next step is prediction where one element is used to predict other elements in the data. Then, the high-pass residual signal, H_i, generated by subtracting the predicted element from the original element, will contain very little energy thereby achieving significant compression. Finally an update step combines residual data from the previous process to reduce the effect of aliasing in low-pass signal, L_i.



(a) (b)

Figure 1: *Wavelet lifting scheme for three-view images. (a) Analysis side. (b) Synthesis side.*

This scheme can be simply adapted to the three-view image system. Its analysis and synthesis diagrams are illustrated in Figure 1 (a) and (b) respectively, where x_0, x_1 and x_2 represent image in left view, middle view and right view, while a_{mn} is the coefficient map from view *m* to view *n*. In traditional filtering such as Haar and 5/3 transform, a_{mn} is fixed to specific values, namely it equals to 0.5 in the predict step and 0.25 in the update step for 5/3 wavelet and equals to 1 in the predict step and 0.5 in the update step for Haar.

If no advanced prediction schemes are exploited in lifting process, i.e. $P(x)$=1 and $U(x)$=1, the low-pass transformed images will suffer from the ghosting artefacts on account of uncompensated shifting of the corresponding data between channels. Furthermore, the high-pass subband will contain considerable energy compromising compression efficiency. This problem might be avoided by enhancing the prediction step within the lifting scheme. Hence, in the context of multi-view image system, the disparity estimation can be usefully applied to the lifting algorithm. This disparity model should be employed without sacrificing the invertible property of lifting transform. The disparity compensated wavelet lifting scheme can be mathematically written as follows.

Defined $D^c_{m\to n}(x)$ is the disparity compensated function predicting for data *x* contained in image view *n* by utilizing the information from image view *m* and generates disparity vectors, $d_{m\to n}$. The low-pass and high-pass components can be written as follows;

$$H_0 = x_1 - a_{01}\cdot D^c_{0\to 1}(x_0) - a_{21}\cdot D^c_{2\to 1}(x_2)$$

$$L_m = x_{2m} + a_{(1)(2m)}\cdot D^c_{1\to 2m}(H_0) \quad m = 0,1$$

The reverse transforms are shown as follows:

$$x'_{2m} = L'_m - a_{(1)(2m)}\cdot \widetilde{D}^c_{1\to 2m}(H'_0, d_{1\to 2m}) \quad m = 0,1$$

$$x'_1 = H'_0 + a_{01}\cdot \widetilde{D}^c_{0\to 1}(x'_0, d_{0\to 1}) + a_{21}\cdot \widetilde{D}^c_{2\to 1}(x'_2, d_{2\to 1})$$

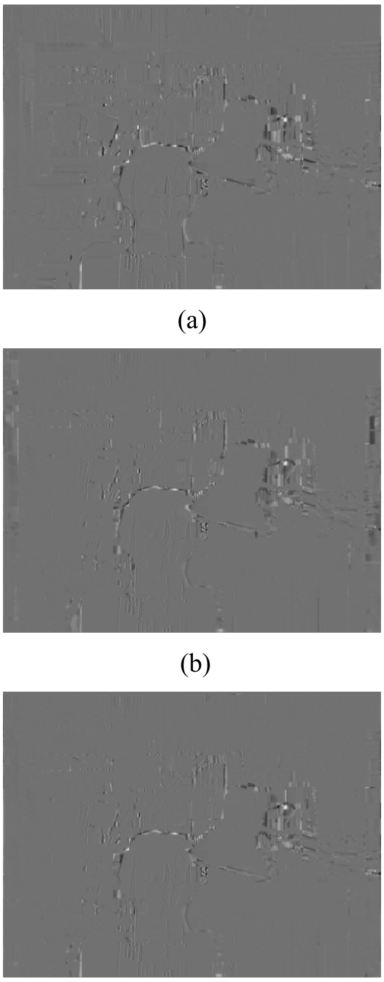
where x' identifies lossy mode of *x* and $\widetilde{D}^c_{m\to n}(x, d_{m\to n})$ represents perfect inverse disparity compensation. Noticeably, in lossless situation, $L_m = L'_m$ and $H_0 = H'_0$, the invertible ability of lifting scheme is preserved.

4. Proposed Scheme

4.1 Hybrid Prediction

Adapting the lifting-based motion transform to multi-view image system in an attempt to remove the redundancy among channel views is not conventional. Unlike the successive frame along a single-view video, the pictures simultaneously captured from multiple cameras are not too similar, since cameras are positioned at various angles to facilitate the 3D scene capture. Therefore, the choice of wavelets filters will be an important parameter in this approach. The longer filters such as the 9/7 biorthogonal wavelet transform are not suitable, since they may lead to poor prediction in occluded areas and also introduce numerous ghosting artefacts. Though, smaller filter such as 5/3 wavelet transform introduces fewer ghosting artefacts, it is still unable to cope with the occlusion absolutely. In contrast, much smaller filters such as the two tap Haar filter will be ideal for dealing with occlusion but are not suitable for achieving high compression. This presents a challenge in terms of an optimal wavelet for multi-view image/video compression. Hence, the hybrid wavelet lifting is proposed as one of the suitable approaches for multi-view image/video coding.

In the hybrid approach, the correlations among consecutive views are efficiently removed by 5/3 wavelet lifting scheme, while the occlusion regions are processed by Haar filters. The occlusion areas in one of the middle view will normally be uncovered/present in



(a)

(b)

(c)

Figure 2: *High-pass images of Head test image. (a) Haar lifting (b) 5/3 wavelet lifting (c) Hybrid lifting*

another side view. Therefore, instead of employing 5/3 transform for the whole image; the occluded areas would be better predicted by using Haar transform in the areas of the image in which they are visible. However, the process for detecting the occlusion has to be combined. To avoid the transmission overhead of additional bits, the local information is exploited to identify which filter is used for each block. However, such local information appearing at encoder and decoder has much opportunity to be different from each other because of lossy conditions, e.g. data stream truncation and channel loss. So, the filter used for a specific block at encoder and decoder may disagree that causes the worse predicting results. On the other hand, side information has to be transmitted or stored to achieve reliable occlusion detection at the encoder and decoder. This later approach is related to the bi-directional prediction in video standards, e.g. MPEG-2, H.264 and etc. where information utilized to indicate the prediction modes is included in bit stream. In the standards, the displacement between the original and predicted values (such as sum of absolute difference (SAD)) is used to determine the most suitable frame for prediction.

In our investigating for the best image compression, the SAD cost after disparity compensation is employed to determine the optimal views for each block. Following the block-based disparity estimation process, two best matching blocks for each block in the middle view are found, one by predicting from left view, P_{ij}^L, and one by predicting from right view, P_{ij}^R. Hence, for each block, the two disparity vectors and SAD are recorded. To get the displacement of 5/3 wavelet scheme, the new SAD cost is merely calculated by summing up the absolute values of the difference between the original image and the average of the existing predicted values, $P_{ij}^{5/3} = 0.5 \cdot (P_{ij}^L + P_{ij}^R)$. It is noticeable that this approach does not require any additional complex computations. The minimum among the three SAD values identifies the type of filters to be used for each block, either Haar with left view, Haar with right view or 5/3 wavelet.

After selecting the optimum filter, therefore, a_{01} and a_{21} matrices do not contain only one specific value but depend whether that block belongs to clear or occluding areas. In the other words, it indicates which filter is used for a given block. For instance, if the block in row i^{th} column j^{th} of middle view, x_1, can be seen in all views, then $a_{01}(i, j)= a_{21}(i, j)= 0.5$ since it is belong to 5/3 wavelet. Alternatively, if it is visible in only one view such as x_0, then $a_{01}(i, j)=1$, $a_{21}(i, j)=0$ and is filtered by

Haar. The high-pass images from block-based disparity compensated Haar, 5/3 wavelet and hybrid lifting scheme with half-pel accuracy are shown in Figure 2. It can be seen that the residual image from hybrid prediction contains less energy than Haar or 5/3 filtering alone.

4.2 Occlusion Detection

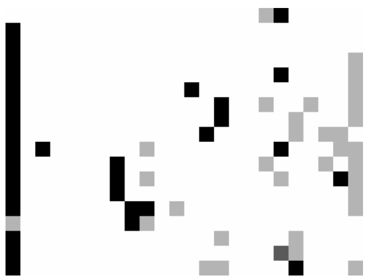
The objective of the previous section is to improve the efficiency of compression by exploiting the hybrid prediction scheme. In this section, the occlusion detection has been proposed in order to gain true disparity, which is needed to accomplish view synthesis in 3D rendering. Basically, the estimated disparity of a pixel is reliable only if a pixel is visible in both considering view and reference view. On the contrary, the defective disparity is possibly generated if that pixel is matched to the improper corresponding pixel. This scenario always takes place if that pixel is occluded in the reference view. Therefore, the effective occlusion detection is necessary to avoid the mismatch areas that would cause the ambiguity of depth information when such disparities are exploited to synthesize or construct a virtual view. Besides, the occlusions cannot be disregarded, since by psychophysical evidence the human visual system exploits occlusion as a positive clue to depth, rather than a hindrance [NS90].

In this paper, we exploit the disparity consistency between two views to detect occlusion areas. This algorithm takes the advantages of the fast and simplest implementation of block-based disparity estimation. It is suitable to compare the performance to that of approach proposed in previous section, because no extra bit overhead is required and the coding algorithm for side information is similar.

The disparity consistency lies on the fact that two views are taken from the same scene, so they are viewing roughly the same. The horizontal disparity vectors derived from left-to-right and right-to-left should be negatives from each other. An occlusion block is detected when the following error is presented:

$$Error_{m\to n}^{(i)(j)} = d_{m\to n}^{(i)(j)} + d_{n\to m}^{(i+d_{m\to n}^{(i)(j)})(j)}$$

The block (i, j) contained in view n is defined to be the occlusion block if $Error_{m\to n}^{(i)(j)} \neq 0$. Therefore, the optimum filters for each block of the middle view are defined by $Error_{0\to 1}$ and $Error_{2\to 1}$. When both $Error_{0\to 1}$ and $Error_{2\to 1}$ equal to zero, it obviously shows that this block appear in all views and its redundant details could be efficiently removed by filtering with 5/3 wavelet lifting scheme. If $Error_{0\to 1} \neq 0$ or $Error_{2\to 1} \neq 0$, it is able to imply that the perfect matching block is not found in Left view or Right view respectively and it should not be predicted by such reference view. This means that the Haar filter could apply with the left view reference when $Error_{0\to 1} = 0$ and $Error_{2\to 1} \neq 0$, whist the right view reference is exploited when $Error_{0\to 1} \neq 0$ and $Error_{2\to 1} = 0$. The optimum filters selected by concerning the best compression following the algorithm proposed in the previous section are illustrated as the filter type map in Figure 3 (a), while Figure 3 (b) shows the selected filter via the disparity consistency algorithm. Note that the white blocks represent the blocks that use 5/3 wavelet, whist the grey blocks and the black blocks show areas where Haar filter is exploited with right view reference and left view reference respectively. The disparity map from blockwise prediction with 16x16 pixels by SAD criteria and disparity consistency criteria are illustrated in Figure 4.



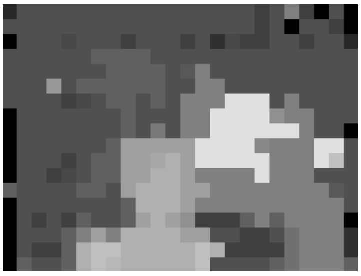
(a)

(b)

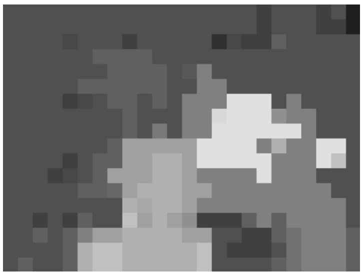
Figure 3: *The filter type: white blocks, grey blocks and black blocks represent the 5/3 filtering, Haar filtering with right view and Haar filtering with left view respectively. (a) SAD criteria (b) Disparity consistency criteria.*

4.3 Adaptive Weighing in Update Step

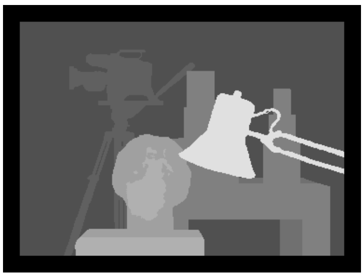
In this section a modification to the update step is proposed. In the update step, the residual from disparity compensated prediction is added to the even images. It is possible to reduce noise and aliasing if and only if the disparity is accurately captured. The erroneous matching tends to introduce ghosting artefacts in low-pass signal. These artefacts appear as the repeat of object boundaries that contain high energy in the high-pass image. In other words, if a block in high-pass image has high energy, it implies that blocks around object boundary will lead to ghosting artefacts.



(a)



(b)



(c)

Figure 4: *The disparity map from block-based estimation (size=16x16) from (a) SAD criteria (b) Disparity consistency criteria. (c) True disparity map*

This visual problem also occurs in lifting based video coding schemes. Various adaptive algorithms for update step have been proposed by using the local information to avoid the extra coded bits [PH02], [MT03]. In the same location of a block, it is not strictly necessary to have the same factor in update step between the encoder and decoder, since it is not used in further prediction. In this paper, we use the normalized energy of the high-pass signal by assuming its relationship directly to ghosting artefacts in low-pass images, i.e. the coefficients of $a_{(1)(2m)}$ are given by normalized energy of $D^{c}_{1\to 2m}(H_0)$ at encoder and $\widetilde{D}^{c}_{1\to 2m}(H'_0, d_{1\to 2m})$ at decoder.

4.4 Overlapped block disparity compensation

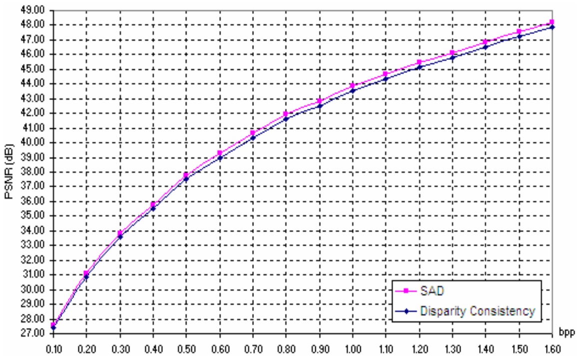
Although, the block-based matching is straightforward to implement, one disparity vector representing the whole pixel in a block may not be sufficient to describe the actual disparity. This tends to produce blocking artifacts which are discontinuities introduced at block borders in the predicted image. They lead to annoying horizontal and vertical edges which are highly visible for the human eye, especially at low bit rates.

To overcome this problem, Overlapped Block Disparity Compensation (OBDC) is employed here to reduce blocking artefacts and disparity compensation error. Moreover, it has the advantages that no change in search range is required and no extra side bits are generated. This is achieved by linearly combining the predictions generated by using multiple disparity vectors, including a block's disparity vector and its neighbours. The OBDC is exactly adapted from the Overlapped Block Motion Compensation (OBMC) [OS94]. The OBDC could be applied to the predict step without obliterating the invertible property of lifting scheme.

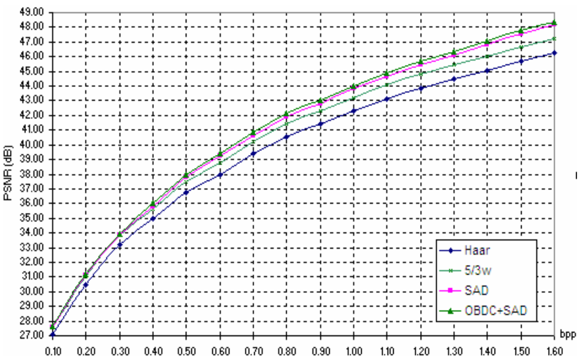
5. Experimental Results and Discussion

In this section, simulation results from the proposed scheme described in section 4 are presented. The hybrid prediction is exploited to increase the accuracy in prediction step. The adaptive weighing factors in update step are identified by normalized energy of high-pass signal in that block, namely $a_{(1)(2m)}$ will be equal to 0.5, 0.25, 0.125 or 0 if such normalized energy is less than 0.25, 0.5, 0.75 and 1 respectively. The performance of this proposed invertible approach was compared to that of the Haar and 5/3 transform with block-based disparity compensation without the OBDC.

The simulations were conducted with the standard multi-view test images, Head. Although the sequences contain five views, three consecutive views were selected for investigating the performance of the proposed three-view image codec and the objective results were averaged. Figure 5 illustrate the average luminance (Y) PSNR of the reconstructed Head images. The chrominance components (U,V) are compressed separately but share the desired target bits. The proposed scheme does not show much improvement in very low bit rates, but shows gains of to 2 dB and 1 dB over Haar and 5/3 wavelet respectively at higher rates. This is because, at low bit rate, most target bits are spent for coding side information, which includes the extra bits used for identifying the choice of filters in the hybrid prediction scheme, instead of being used for coding texture of image. The PSNR results are slightly decreased when the occlusion detection by using disparity consistency information are applied. The results outperform than exploiting only Haar filter or only 5/3 wavelet transform. However, the block-based algorithm assumes one depth in each block that produces some errors. The enhance algorithms should



(a)



(b)

Figure 5: *The results of proposed codec. (a) Comparing between SAD criteria and disparity consistency criteria. (b) Comparing to Haar and 5/3 wavelet scheme.*

be applied to disparity estimation and occlusion detection, such as adaptive block size and occlusion constraint that detects the disparity jump at the occluding boundary [EW02]. Figure 6 show the subjective results from SAD and disparity consistency criteria when coding at 0.6 bpp. Although, the PSNR of the second approach is less than the first one by 0.3 dB, the reconstructed images could not be easily noticed the different quality.

Subsequently, the OBDC with weighting determined by a Raise Cosine window is implemented in the prediction step. The objective results are included in Figure 5 (b). The subjective quality of reconstruction is significantly improved.

(a)

(b)

Figure 6: *The reconstructed middle images of Head (a) SAD criteria (b) Disparity consistency criteria.*

6. Conclusions

A novel multi-view image coding based on lifting has been proposed. By exploiting the wavelet lifting scheme with block-based disparity compensation, improved coding performance is reported. A hybrid prediction was proposed to cope with occlusions which are often present in multi-view images. We proposed a simple computationally efficient occlusion detection and view selection scheme comparing between the SAD criteria and disparity consistency. Furthermore, an adaptive weighing in update step is included to reduce ghosting artifacts. An overlapped block disparity compensation approach is also proposed to reduce blocking artifacts. The results show that proposed approach can lead to significant improvements in both objective and subjective quality of the reconstructed image. It is clear that the lifting based approach has a number of advantages in multi-view coding including gain improvements of up to 1.5 dB and 1.2 dB over a conventional Haar and 5/3 transform block-matching codec respectively. In future, the codec will be extended to a scalable multi-view image and video system.

References

[DGR*88] I.Dinstein, G.Guy, J.Rabany, J.Tzelgov, A.Henik, On Stereo Image Coding, *IEEE Proc. ICPR* (Nov 1988), pp. 357-359.

[EW02] G. Egnal and R. P. Wildes, Detecting Binocular Half-Occlusions:Empirical Comparisons of Five Approaches, *IEEE Trans on Pattern Analysis and Machine Intelligence* (Aug 2002), **24**(8): 1127-1133.

[LLL*01] L. Luo, J. Li, S. Li, Z. Zhuang and Y. Zhang, Motion Compensated Lifting Wavelet and Its Application in Video Coding, *IEEE Proc. ICME* (Aug 2001), pp. 481-484.

[LT98] R.A.Lane and N.A.Thacker, Tutorial: Overview of Stereo Matching Research, Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester (1998).

[Luk96] Lukacs, M.E., Predictive Coding of Multi-viewpoint Image Sets, Proceedings of the 1986 IEEE-IECEJ-ASJ International Conference on Acoustics, Speech, and Signal Processing (1996), Tokyo, Japan, pp. 521-524.

[MSJ96] Jeffrey S. McVeighaa, M.W. Siegel b, A.G. Jordana, Intermediate view synthesis considering occluded and ambiguously referenced image regions, *Signal Processing: Image Communication 9* (1996), pp. 21-28.

[MT03] N. Mehrsereht, D. Taubman, Adaptively Weighted Update Steps in Motion Compensated Lifing Based Scalable Video Compression, *IEEE proc. ICIP* (Sept 2003), **3**: 771-774.

[NS90] K. Nalayama and S. Shimojo, Da Vinci Stereopsis: Depth and Subjective Occluding Contours from Unpaired Image Points, *Vison Res.*(1990), **30**(11):1811-1825.

[Ohm94] J. Ohm, Three dimensional subband coding with motion compensation, *IEEE Trans. on Image Processing* (Sep 1994), **3**: 559-571.

[OS94] M.T. Orchard, G.J. Sullivan, Overlapped block motion compensation: an estimation-theoretic approach, *IEEE Trans. Image Processing* (Sept 1994), **3**(5): 693 – 699.

[PB01] B. Pesquet-Popescu and V. Bottreau, Three-dimensional Lifting Schemes for Motion Compensated Video Compression, *IEEE Proc. ICASSP'01* (May 2001), pp. 1793-1796.

[Per92] M.G. Perkins, Data Compression of Stereopairs, IEEE Trans. Communications (1992), **40**(4): 684-696.

[PH02] G. Piella and H.J.A.M. Heijmans, Adaptive Lifting Schemes With Perfect Reconstruction, *IEEE Trans. Signal Processing* (Jul 2002), **50**(7): 1620-1630.

[ST03] A. Secker and D. Taubman, Lifting-Based Invertible Motion Adaptive Transform (LIMAT) Framework for Highly Scalable Video Compression, *IEEE Trans. on Image Processing* (Dec 2003), **12**(12): 1530-1542.

[Swe98] W. Swelden, The Lifting Scheme: A Construction of Second Generation Wavelets, *SIAM Journal of Mathematical Analysis* (Mar 1998), **29**(2): 511-546.

[ZFH*98] M. Ziegler, L. Falkenhagen, R. Horst, D. Kalivas, Evolution of Stereoscopic and 3D video, *Signal Processing: Image Communication 14* (1998), pp. 173-194.