Link to published version (if available):
10.1109/TCSVT.2009.2031371

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

In-Band Disparity Compensation for Multi-view Image Compression and View Synthesis

N. Anantrasirichai, *Member, IEEE*, C. Nishan Canagarajah, *Member, IEEE*, David W. Redmill and David R. Bull, *Member, IEEE*

Abstract—**This paper presents a novel framework to achieve scalable multi-view image compression and view synthesis. The open-loop wavelet lifting scheme for geometric filtering has been exploited to achieve SNR scalability and view-type scalability (mono-, stereo-, or multi-view). Spatial scalability is achieved by employing in-band prediction which removes correlations among subbands (level-by-level) via shift-invariant references obtained by Overcomplete Discrete Wavelet Transforms (ODWT). We propose a novel In-band Disparity Compensated View Filtering (I-DCVF) approach, akin to the Motion Compensated Temporal Filtering (MCTF), for achieving a scalable multi-view codec. In our codec, hybrid prediction is proposed to deal with occlusions, and a novel cost function in dynamic programming (DP) for disparity estimation is introduced to improve view synthesis quality. Experiments show comparable results at full resolution and significant improvements at coarser resolutions compared to a conventional spatial prediction scheme. View synthesis efficiency is extensively improved by utilizing disparity estimation from the proposed DP approach.**

Index Terms—**Multi-view image, disparity compensation, occlusion, overcomplete discrete wavelet transform, dynamic programming, scalability.**

I. INTRODUCTION

VISUAL technologies, especially three-dimensional (3D) imaging systems, have found many applications in remote surveillance, medical imaging, telerobotics, entertainment and virtual reality. 3D images/videos are typically constructed from a series of multi-view sequences captured from multiple cameras located at different positions. This results in very high raw data rates, or storage space compared to monoscopic systems. Fortunately, the clear context of one view appearing as shift position from other views, which is known as disparity, leads to significant redundancies in such sequences. This geometric correlation can be removed by efficient prediction and compression algorithms. For example, in [1], the furthest

The authors are with the Department of Electrical and Electronic Engineering, University of Bristol, Bristol, BS8 1UB, UK (e-mail: n.anantrasirichai@bris.ac.uk;nishan.canagarajah@bris.ac.uk; david.redmill@brist.ac.uk; dave.bull@bris.ac.uk.

right view is independently coded and the remaining views are predicted using this reconstructed view. Due to the finite viewing and occlusion area, this method may perform poorly compared to the independent coding. Moreover, propagation of transmission errors can result from the recursive decoder structure exploited in the first view coding. Another approach [2] constructs intermediate views by utilizing the disparity information from the left and the right views. However, the errors from disparity estimation can be exacerbated after the interpolation. In this paper, we propose an efficient hybrid prediction algorithm to deal with the occlusions and yield better performance.

Multi-view systems require both high compression, as well as a variety of other features to support heterogeneous systems and networks. Multi-view image coding should provide scalability mechanisms similar to those provided by scalable video codecs. Scalable image/video encoding allows extraction of specific bitstream subsets with different transport and presentation properties. Recently, the discrete wavelet transform (DWT) has become an attractive choice to meet these requirements. In this work, we extend the wavelet scheme for image/video coding to remove spatial, temporal and geometric redundancies present in multi-view sequences. To exploit temporal information, a method based on Motion Compensated Temporal Filtering (MCTF) has been introduced [3-6], which inherently facilitates the spatial-temporal-SNR scalability. However, these systems typically ignore the possibility of occlusion which is common place in multi-view image system. This problem can be solved with our proposed hybrid prediction scheme [7] and enhanced disparity estimation using dynamic programming [8].

Based on the wavelet lifting scheme, MCTF contains flexibility in the number of decomposition levels and the choice of filters. Moreover, it operates as open-loop prediction, so temporal drift problem is eliminated and error resilience is improved. However, reconstructed images for the limited display size still suffer from the drift problem. Direct prediction at the smaller size could eliminate this difficulty by using in-band motion estimation which predicts subband-to-subband [9]. However, the shift variance as a result of the decimation process in the wavelet transform frequently causes the inefficient estimation in high-pass subbands. This problem can be avoided by omitting the subsampling process after filtering leading to an Overcomplete Discrete Wavelet
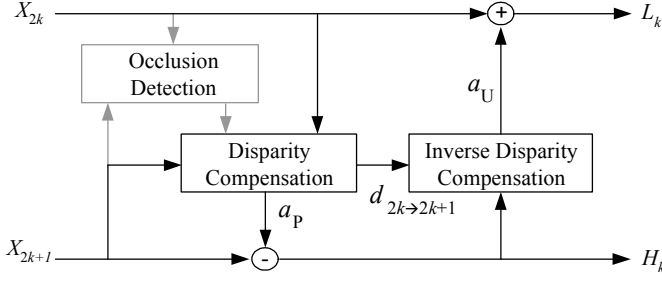
Fig. 1. The forward lifting scheme for multi-view image.

Transforms (ODWT) [10-13]. In this paper, a novel in-band disparity compensated view filtering (I-DCVF) is proposed for multi-view image compression. The proposed in-band prediction can remove the redundancy between similar subbands in different views. It improves the decoding efficiency at smaller spatial resolutions. Moreover, blocking artefacts are eliminated by using the in-band prediction, since the boundaries are filtered out during inverse DWT [14].

View synthesis is a key requirement in multi-view systems. This feature needs reliable disparity information, hence disparity estimation becomes an important element for 3D visualization. Due to occlusions, imperfect camera calibrations, and imperfect light balance, accurate disparity information is hard to achieve. An acceptable compromise is to ensure viewers feel the different depth of each object and also feel comfortable. In this paper, we use the DP to estimate the global optimum disparity/depth fields. The DP for estimating the disparity between stereo images was first been introduced in [15] and it shows better performance over traditional matching schemes. Moreover, the DP can deal the proper disparity estimation and the occlusion detection simultaneously. A novel cost function proposed in [8] is used for searching the minimum-cost path, and then the performance of view synthesis at numerous spatial resolutions is investigated.

In this paper, we present a compression/synthesis framework to achieve scalable multi-view image coding with the following novel contributions; i) a hybrid prediction scheme to deal with the presence of occlusions, ii) a novel cost function in the DP for disparity estimation in order to obtain depth and occlusion information simultaneously thereby improving view synthesis, iii) an in-band DCVF to achieve spatial, SNR and view-type scalabilities (mono-, stereo-, or multi-view), and iv) a combination of disparity map and low-pass subband to reach an acceptable compromise between compression and view synthesis performances.

The rest of the paper is organized as follows: section II briefly explains the fundamentals of the wavelet lifting scheme and its adaptation to multi-view environments. The proposed codec is described in section III. Subsequently, in section IV, the coding system combines the disparity map obtained by using the DP in order to improve view synthesis. The experimental results are presented in section V followed by conclusions and future work in section VI.

II. GEOMETRIC LIFTING SCHEME

The *lifting scheme* introduced by Sweldens [16-17] is a technique for constructing or factoring wavelet filters into basic building blocks. A related structure called ladder networks was proposed earlier by Bruekers [18]. Basically, the forward wavelet lifting method decomposes the wavelet transforms into a set of stages. The operation starts with a split step, which divides the data set into groups, normally composed of odd and even samples. The next step is prediction where one group is used to predict other group. Then, the high-pass residual signal, H_k, generated by subtracting the predicted element from the original element, will contain very little energy thereby achieving significant compression. Subsequently, an update step combines residual data from the previous process to reduce the effect of aliasing in low-pass signal, L_k. Obviously, in a wavelet tree, H_k should be encoded more coarsely and can be dropped at a receiver if bit rate is constrained. In this manner, the output bitstream of the proposed encoder achieves (SNR, temporal) scalability.

A. MCTF to DCVF

In video coding, the wavelet transform is applied to filter data in the intra frame and also along the temporal direction, called three-dimensional discrete wavelet transform (3D-DWT). The performance of prediction and update steps in the temporal direction can be improved by utilizing motion compensation leading to the motion compensated temporal filtering (MCTF) algorithm. Ohm has proposed 2-tap Haar filter for MCTF by using the block-based motion compensation and showed that a non invertible transform will affect coding gain if there are numerous disconnected pixels between blocks [19]. In [4], a non-linear lifting framework for temporal wavelet transform has been introduced which is shown to be superior to the conventional temporal filtering. Luo et al. have employed a bi-directional motion estimation in each lifting step of temporal direction [5], while the longer filters have been introduced with a higher coding gain for video compression by Golwelkar and Woods [20]. Secker and Taubman have developed the highly scalable video coding approach [6]. It is based on the lifting with adaptive motion compensation. However, they have ignored the possibility of occlusion which is common problem in multi-view imaging systems.

The use of lifting scheme in view direction without advanced prediction scheme will result in the ghosting artefacts on account of uncompensated shift of corresponding data between views. Furthermore, the high-pass subband will contain considerable energy thereby compromising compression efficiency. In the context of multi-view image system, disparity estimation can be usefully applied to the lifting algorithm. Obviously, the view-type scalability is supported by this geometric wavelet transform [7,21-23]. A number of views are selectively determined at the receiver according to the type of display modes, mono-, stereo-, or
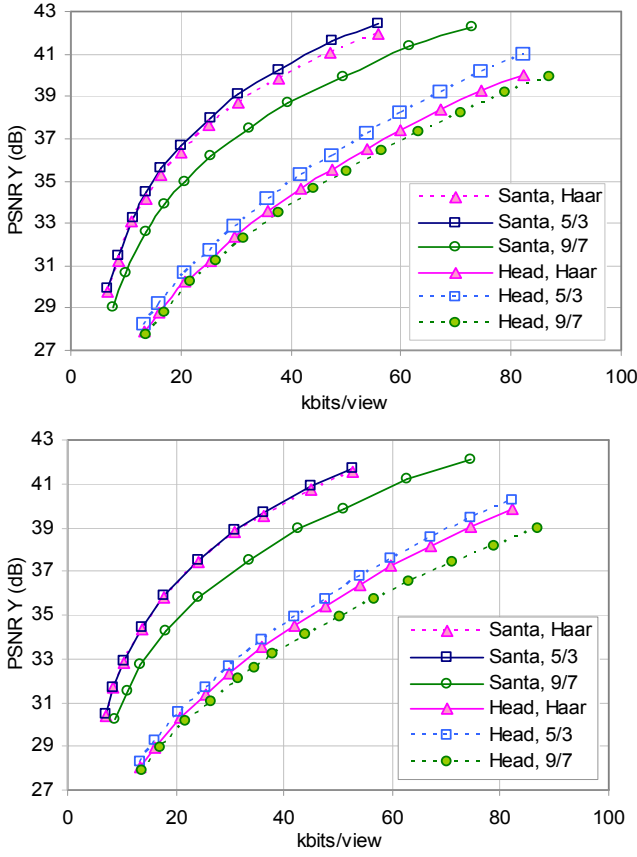
Fig. 2. The average PSNR quality of the different wavelet decomposition. Top: narrow-baseline geometry. Below: wide-baseline geometry.

multi-views. Then, only the selected views are decoded and displayed.

Errors in disparity information come from two sources; the presence of noise in images and the invisibility of corresponding points in reference images. The latter case is dependent on the location of cameras. Three or more cameras are often able to compensate such areas and the disparity compensated function is then modified to include the known occlusion information. The forward transform of the disparity compensated view filtering (DCVF) are displayed in Fig. 1. The disparity compensation and the inverse disparity compensation are utilized to improve the performance of prediction and update step respectively. The occlusion detection is also used to reduce the disparity mismatch error. In section IV, we propose disparity estimation based on the DP that simultaneously detects occlusions.

B. Types of Wavelet Filter for Multi-view Imaging

Adapting the wavelet lifting scheme to multi-view differs from MCTF with single-view video. The pictures simultaneously captured from multiple cameras are usually less similar. Therefore, the choice of wavelets filters will be an important parameter in geometric view filtering.

We have investigated three filters: Haar (short-length filter), 5/3 wavelet (medium-length filter) and 9/7 wavelet (long-length filter), for two types of multi-view sequences: narrow-

baseline and wide-baseline geometries. The simulations use conventional block-based disparity estimation with 8x8 block size to multi-view test sequences, Head (five views with size 384x288) and Santa (nine views with size 320x240). The 5-view Head sequence is extended symmetrically at the furthest left and right view until it contains totally nine views when the 9/7 wavelet transform is applied. For the wide-baseline geometry, only odd views or non-consecutive views are used. This increases the presence of occlusions. The performance of each filter type is illustrated in Fig. 2. The 9/7 wavelet gives significantly lower performance for multi-view sequences with either narrow-baseline or wide-baseline geometry. It utilizes information from nine views that generates a series of ghosting artefacts around object boundaries. Moreover, the disparity estimation error could be exacerbated after many steps of lifting. Although the 5/3 wavelet outperform other filters, the results still contain errors around occlusion areas which are not visible in both reference views. Better results in such occlusion areas are predicted with only the visible view reference. For the wide-baseline test which presents more occlusions, the Haar filter yields performance close to that of 5/3 wavelet. This suggests that the Haar filter deals well with the occlusion problems.

III. MULTI-VIEW IMAGE CODEC WITH IN-BAND DISPARITY COMPENSATED VIEW FILTERING

The proposed I-DCVF scheme has been developed from the disparity compensated lifting scheme, with the prediction step is performed after the spatial transformation. The direct subband-to-subband prediction makes disparity estimation at coarse resolution more precise. Additionally, the proposed hybrid prediction scheme enhances the disparity estimation, since it can compensate for the effect of the occlusion [24].

Our proposed geometric wavelet lifting scheme with in-band hybrid prediction is illustrated in Fig. 3. The encoder starts the operation by performing spatial wavelet transform up to R levels to each view and gaining subbands s, where $s = \{\mathcal{L}^R, \mathcal{H}^u\}$, $\mathcal{H} = \{HL, LH, HH\}$, $\mathcal{L} = \{LL\}$ and $1 \leq u \leq R$. Then, the transformed images, X_v^s, are filtered along the geometric direction. The wavelet coefficients of the reference views are converted to ODWT coefficients by using the *Complete-to-overcomplete discrete wavelet transform* (CODWT) method [25-26]. The *DE/DC* enhances the prediction step in the lifting scheme by employing both ODWT coefficients $X_{2k}^{os} = \mathbf{w}_{ODWT,2k}^u$ of $(2k)^{\text{th}}$ reference images, and $X_{2k+2}^{os} = \mathbf{w}_{ODWT,2k+2}^u$ of $(2k+2)^{\text{th}}$ reference images, to predict the intermediate view X_{2k+1}^s. It produces the predicted subband, $P_{2(k+i) \rightarrow 2k+1}^s$ and disparity vectors $d_{2(k+i) \rightarrow 2k+1}^s$.

The *filter selector* is applied to provide the hybrid prediction on a block wise basis to generate the coefficients $a_{2(k+i) \rightarrow 2k+1}^s$ depending upon the chosen filter f_{2k+1}^s. However, because of non ideal filters and disparity failure, the coding of
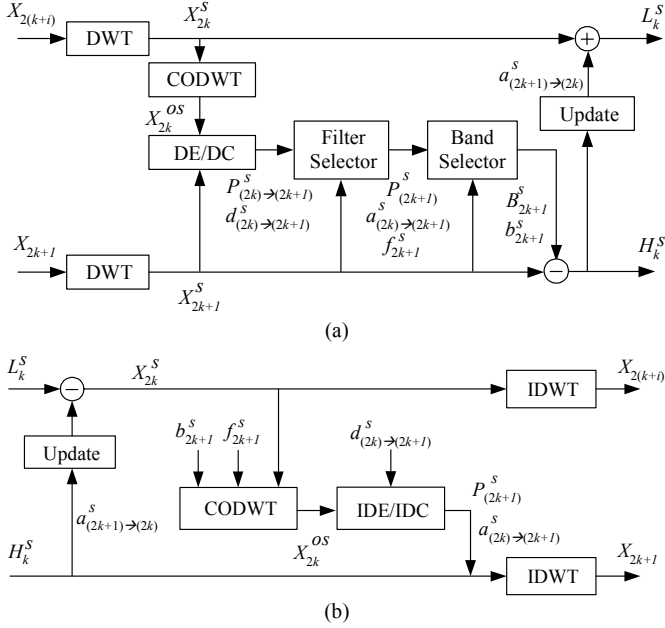
(a)

(b)

Fig. 3. The in-band DCVF of multi-view image system. (a) forward part. (b) reverse part.

TABLE I
PERFORMANCE OF VARIOUS PREDICTION SCHEMES

MSE of geometric high-pass images				
Sequence	Haar	5/3	Hybrid (SAD)	Hybrid (occlusion)
Santa	11.56	9.31	7.19	9.00
Head	18.14	16.72	10.34	12.85
MSE of disparity maps				
Sequence	Haar	5/3	Hybrid (SAD)	Hybrid (occlusion)
Santa	7.32	5.54	5.14	4.81
Head	4.24	3.14	3.19	2.98

IDWT.

A. In-band Disparity Compensated View Filtering

In the I-DCVF process, each view of multi-view images is iteratively spatially decomposed into various subbands. Then, the geometric wavelet lifting scheme with the disparity estimation and compensation is applied independently to each subband.

After constructing ODWT, each resolution level u will be combined with the critically-sampled subbands $LH^u_{(i,j)}, HL^u_{(i,j)}, HH^u_{(i,j)}$, $0 \le i,j < 2^u$ and LL at coarsest resolution level. The full search of the disparity estimation can be run in a level-by-level fashion. The best matching is found by jointly minimizing the distortion for each triplet of blocks from the $LH^u_{(0,0)}$, $HL^u_{(0,0)}, HH^u_{(0,0)}$ of the current view that correspond to the same spatial domain location of a triplet of blocks from $LH^u_{(i,j)}$, $HL^u_{(i,j)}$, $HH^u_{(i,j)}$ of the reference views. For the coarsest resolution level (containing only LL subband), the estimation process is done separately. The block size of the coarser resolution level will be smaller than that of the successive finer resolution level by factor of two. This achieves better prediction compared to the conventional spatial estimation and requires equal or less DV.

The proposed DCVF provides a compact scalable multi-view image compression across multiple spatial resolutions. Users are allowed to exploit different filter and estimation parameters for each resolution level. Therefore, the disparity maps can be exploited in order to improve view synthesis, which will be discussed in section IV.

B. Hybrid Prediction Scheme

From section II.B, we know that the longer filters are not suitable for multi-view system, since they may lead to poor prediction in occluded areas and also introduce numerous ghosting artefacts. Although, smaller filter gives fewer ghosting artefacts, it is still unable to cope with the occlusion. In contrast, much smaller filters such as the two tap Haar will be ideal for dealing with occlusion but they are not suitable for achieving high compression [20]. This presents a challenge in terms of an optimal wavelet for the multi-view image/video compression. Hence, hybrid wavelet lifting is proposed here as a suitable approach for multi-view image/video coding.

In the hybrid approach, correlations among consecutive

both the residual and the corresponding DV is not necessarily superior to coding of the wavelet coefficients themselves. This mainly applies to the subbands of high frequency, since the residual results of high-pass prediction sometimes have higher energy than the original high-pass subbands. Hence, the *band selector* is applied to choose between the predicted high-pass and the original high-pass subbands. If the original high-pass subband contains less energy, B^s_{2k+1} will be a zero matrix; otherwise, it would be equal to P^s_{2k+1}.

The even views are updated subsequently to obtain the low-pass images L^s_k. The function of the update block is to combine the results of DE/DC that are estimated with information of intermediate view. This consequence should be the inverse direction of $d^s_{2k+1 \to 2k}$ for updating left view or $d^s_{2k+1 \to 2k+2}$ for updating right view. If there is occlusion, the best match in the update step is not equal to the inversion of such DVs. Unfortunately, the erroneous matching tends to introduce ghosting artefacts in the low-pass signal. Therefore, an adaptive weighing scheme is exploited to suppress the coefficients in a block that contains high energy [27-28]. The subbands of the coarsest resolution ($\mathcal{L}^R, \mathcal{H}^R$) are further decomposed, and then coded by using Embedded Block Coding with Optimized Truncation (EBCOT) [29].

The decoder operates as a mirror of the encoder. By subband-to-subband, the odd views are decoded by subtracting the high-pass signal. Then, the ODWT and the prediction coefficients of the selected bands and filters are generated. Using disparity vectors, the predicted signal of the even views is reconstructed via the inverse DE/DC process. The high-pass signal is then added to generate the even views. Ultimately, all reconstructed subbands are combined together and using the
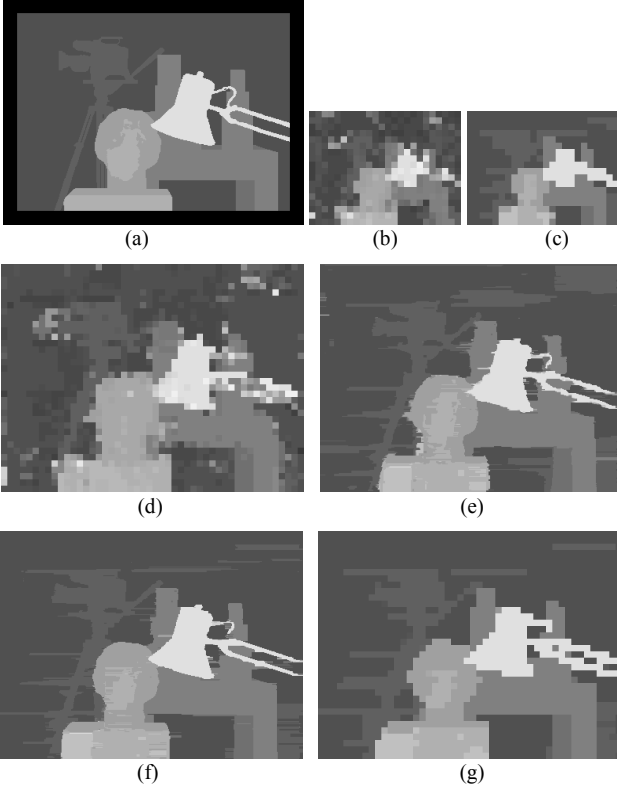
Fig. 4. The disparity maps. (a) The true disparity. The estimated disparity maps at level 2. (a) block-based estimation (4x4) and (b) the proposed DP. The estimated disparity maps at full resolution by (d) block-based estimation (8x8), (e) the proposed DP with 1D camera configuration and (f) the proposed DP with 2D camera configuration. (g) The discrete disparity map of (f) with 8x8 size equivalent to (b).
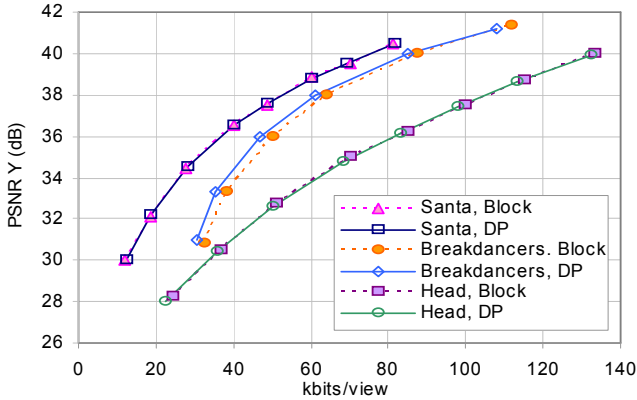

Fig. 5. Rate-distortion performance of the in-band prediction by block-based estimation against the DP.

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 intermediate views will normally be uncovered/present in 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. However, the process for detecting the occlusion has to be defined and side information has to be transmitted or stored to achieve reliable occlusion detection at the encoder and decoder.

The mean square error (MSE) results of three standard test multi-view images are tabulated in TABLE I to demonstrate the results of the hybrid lifting scheme compared to the traditional Haar and 5/3 wavelet transforms by applying the 8x8 block-based disparity estimation. The SAD cost after disparity compensation is employed to determine the optimal views for occlusion compensation. Note that using the SAD criteria basically gives the good compression in terms of objective results with low complexity. Following the block-based disparity estimation process, two best matching blocks for block (i, j) in the odd view are found, one by forward predicting from $(2k)^{\text{th}}$ even view, $P_{2k \to 2k+1}^{ij}$, and another one by backward predicting from $(2k+2)^{\text{th}}$ view, $P_{2k+2 \to 2k+1}^{ij}$. 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. The minimum among the three SAD values identifies the type of filters to be used for each block, either Haar with $(2k)^{\text{th}}$ view, Haar with $(2k+2)^{\text{th}}$ view or 5/3 wavelet. Note that the MSE can be used as criterion to find the best matching, but it produces higher complexity [30].

The hybrid prediction scheme discussed above provides the highest compression efficiency in terms of PSNR values. However, view synthesis algorithms in 3D rendering have been proposed to generate the intermediate views. For good synthesis, the true disparity is needed. The disparity estimation is always reliable if and only if that pixel can be seen in both current view and reference view. The occlusion cannot be disregarded, since by psychophysical evidence the human visual system exploits it to get depth information [31]. The proposed occlusion detection is embedded in DP which will be described in section IV. The result of the hybrid prediction with this occlusion knowledge is also shown in TABLE I.

The hybrid prediction scheme achieves best results over the Haar and 5/3 wavelet with both SAD cost and the estimated occlusion. The error in the residual image of the hybrid prediction with the knowledge of occlusions is a bit higher than that of the SAD cost. However, the estimated disparity map is improved. Note that the original data show the occlusion areas of 6% and 10% for the Head and Santa images respectively, while the hybrid prediction scheme gives the occlusion results of 11% and 16% for the Head and Santa when the SAD cost is used. The proposed prediction schemes always give more detected occlusion areas because of an assumption of homogeneous characteristics over a block. The proposed prediction scheme shows 10% and 15% misused-filter blocks for Head and Santa images respectively. The Santa images contain fewer details in a block compared to those of the Head images. As a result, the proposed prediction scheme, based on a block matching algorithm, shows higher possibility of misusing the appropriate filter. Note that we define that the blocks with occlusion area more than 10% of
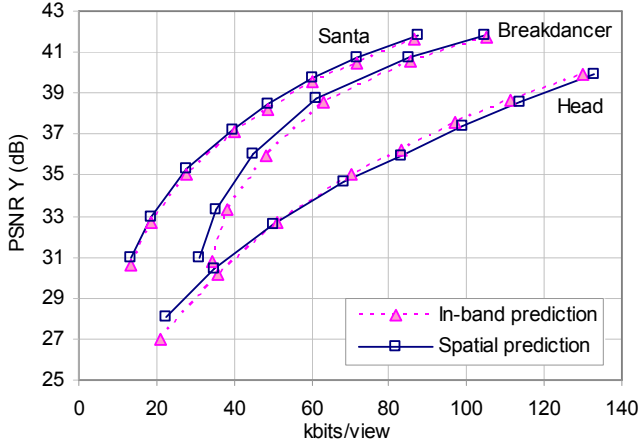
Fig. 6. Rate-distortion performance of the spatial prediction and in-band prediction.
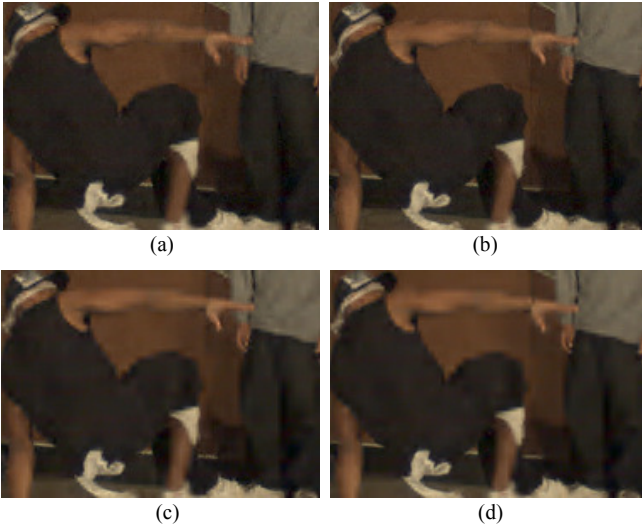

Fig. 7. The full-resolution reconstructed middle views of *Breakdancers* at 60kbits/view (a) block-based spatial prediction (PSNR Y=38.12). (b) DP spatial prediction   (PSNR Y=38.54). (c) block-based 3-level in-band prediction (PSNR Y=37.23). (d) DP 3-level in-band prediction (PSNR Y=37.82).
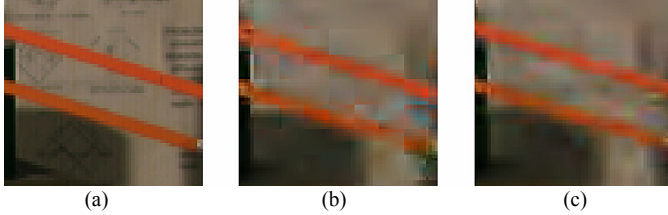

Fig. 8.   The enlarge images at 40 kbits/view. (a) The original image. (b) Spatial prediction. (c) In-band prediction.

the whole block are the free blocks which can exploit either Haar or 5/3 wavelet filters and are not included in the misused-filter blocks.

IV. DISPARITY ESTIMATION FOR LOW-PASS SUBBANDS

As well as providing good image quality for limited bandwidth, an efficient multi-view compression algorithm should also facilitate good view synthesis. To support the virtual view construction, disparity or depth information of the interested objects needs to be stored or transmitted with the texture information. Therefore, conventional disparity estimation, e.g. block-based matching by SAD criteria, may not provide the correct disparity, thereby generating texture and position errors in the synthesised intermediate views.

In this paper, we propose novel cost functions used in the DP to estimate disparity/depth of multi-views. The main advantage of the DP compared to traditional matching schemes is the absence of blocking artefacts or noisy depth maps. Moreover, the DP is an efficient tool for solving multi-stage problems, which enables combined disparity estimation and occlusion detection. The choice of a good cost functions for searching the minimum-cost path is a key aspect of the DP approach. The simplest cost function utilizes the similarity of luminance between two views and the cost of occlusions is constant. I. Cox has proposed the matching process using individual pixel intensity [32]. Although cohesivity constraints are used to deal with the inter-scanlines disparity discontinuities, the ambiguity from imperfect light balance might affect homogenous areas. We therefore exploit the window-based correlation with adaptive window size and shiftable windows in this paper. The accuracy of the occlusion detection can be enhanced by Bayesian method but a probability of occlusion is required [33-34]. N. Grammalidis and M. G. Strinzis have proposed the disparity estimation and the occlusion detection algorithm for multi-view system, but cost defined to identify occlusion is fixed [35]. In this paper, we propose a simple but effective cost function by exploiting confidence information from other cameras. The errors from the matching blocks are employed to calculate the error function. A ceiling is imposed on the error function so that very large individual errors will not dominate the accumulated cost used for defining the global path in DP approach. The original contributions here also include the algorithms for multi-view image extension and for combining horizontal and vertical scanning.

A. Disparity Estimation by Dynamic Programming

The matching pixels along each scanline are searched via the DP with a novel cost functions proposed by the authors of [8]. The key idea comes from the three possible disparity values of each particular pixel, which are equal to, more or less than that of the consecutive pixel. The first case usually occurs in the non-occlusion areas, while the last two cases possibly occur in the occlusion areas. Hence, three costs are defined to each node (i, j) in the DP. C_1 and C_2 are the occlusion costs of the pixel, which are invisible in left view and right view respectively, and C_3 is the cost of the pixel in non-occlusion areas. These costs are expressed as follows:

$$C_1 = \left(\frac{1}{\lambda} - E_l\right) + E_r \,, \ C_2 = \left(\frac{1}{\lambda} - E_r\right) + E_l \,, \ C_3 = E_l + E_r \tag{1}$$

$$E_l = \frac{e_{ij}^2(d)}{1 + \lambda \cdot e_{ij}^2(d)} \,, \ E_r = \frac{e_{ij}^2(\frac{B^{R_1}}{B^{L_1}} d)}{1 + \lambda \cdot e_{ij}^2(\frac{B^{R_1}}{B^{L_1}} d)} \tag{2}$$

where $e_{ij}^2(k) = \sum\limits_{(i,j) \in w_{ij}} a_{\tilde{i}\tilde{j}} \cdot (p_{ij}^C - p_{(i)(j+k)}^R)^2$, $a_{\tilde{i}\tilde{j}}$ is a weighing coefficient that is inversely proportional to the distance from the pixel (*i,j*), and p_{ij} is the intensity of pixel (*i,j*). The B^{L_1} and B^{R_1} are the baseline between the current camera and the closest left and the closest right cameras respectively, while λ is the occlusion parameter.

Finally, the least accumulated cost of the last pixel in a scanline is selected to identify the optimum path. After tracking back along the optimum path, the estimated disparities are generated and the occlusion areas are simultaneously marked at the pixels where the occlusion cost dominates.

The incorrect path might appear at the occlusion areas composed of similar details of the neighbouring non-occlusion areas. However, if more cameras are available, this problem could be eliminated by using information from other reference views as following subsections.

1) Multi-view Image Extension: To extend the three-view disparity estimation to the general multi-view disparity estimation for a linear configuration, the error from all the reference views are compared and the minimum is selected. Better prediction is achieved by exploiting the information from other view references. However, these error functions are adapted to the cost C_3 only, whilst the cost C_1 and C_2 are still calculated from the closest left and the closest right views. This is because the smaller error might lead the mistaken path in occlusion areas, i.e. C_3 would dominate instead of C_1 or C_2.

2) Combining Horizontal and Vertical Scanning: The proposed scheme is extended to the planar camera configuration. The estimated disparity from scanning in one camera axis can be used to modify the cost for each node of the scanning process in another camera axis by pre-marking the occlusion areas. After vertical scanning, for example, the possible occlusion regions of the horizontal scanning are marked from the vertically estimated disparity map, and then the costs C_1 and C_2 of such regions are modified with the proportion α as follows: $C_1^{new} = \alpha \cdot C_1$ and $C_2^{new} = \alpha \cdot C_2$, where 0 < α < 1. As a result, the disparity and the occlusion in this horizontal scanning are more reliable than the one that does not exploit information from the vertical scanning. Moreover, the result of the vertical scanning where C_3 dominates is subsequently employed to replace the result of the horizontal scanning which is marked as the occlusions.

The results of the proposed disparity estimation by DP are shown in Fig. 4. For the Head sequence, the percentage of the pixels whose absolute disparity error are greater than 1 is 4.05% which is slightly better than the DP proposed in [36],

but it is worse than some algorithms as a comparison in [37] (for Santa sequence, the proposed scheme achieves 3.77% error, which the DP in [36] achieves 3.85%). However, our proposed algorithm is obviously simpler by exploiting only intensity constraint. Moreover, the results of the estimation are good enough to use with proposed I-DCVF schemes thereby achieving both compression and view synthesis performance which will be discussed in the following sections.

B. Low-Pass Subband Prediction Improvement

As discussed in the previous section, the low-pass subband prediction does not suffer from the decimation process in wavelet decomposition. Therefore, the disparity estimation for low-pass subband can be directly applied without CODWT. However, the too small size of image might diminish the estimation performance. We suggest operating the disparity estimation at the full resolution, and then the disparity results are sub-sampling into the proper size associating with the resolution level. Afterward, we make it discrete associating with the block size by using median value of disparities within a particular block. However, if the frequency of the maximum value is greater than one quarter of a number of pixels in a block, this value is selected thereby leading the clear and sharp edges of front objects. The estimated disparity map is transformed to discrete version in order to reduce the number of disparity vectors. Note that the high-pass subband estimation is still operated as explained in section III.A. The comparison between the block-based estimation and the DP in Fig. 5 shows that the estimation by using the DP does not deteriorate the quality of reconstructed images. The estimated disparity maps of the Head sequence are improved as displayed in Fig. 4 (b)-(c) and (d)-(g) for the quarter resolution and the full resolution respectively. The estimated disparity maps of the conventional block-based matching present noisy disparities in homogenous areas and also the errors because of unfit block size with the position of object boundary. Interestingly, the results of Breakdancers show the improvement of the DP at low bitrate. This is because the estimated depth of the DP is significantly smoother than that of the block-based estimation thereby achieving lossless compression.

V. EXPERIMENTAL RESULTS

This section presents simulation results for the proposed scheme. Hybrid prediction is exploited to increase the accuracy in prediction step. Adaptive weighing factors in update step are identified by normalized energy of high-pass signal in that block, namely $a_{2k+1 \rightarrow 2(k+i)}^s$ 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 simulations were conducted with three standard multi-view test images; Breakdancers, Head, and Santa, containing eight, five, and nine views respectively. Non-consecutive views were selected for investigating the performance of the proposed image codec, while the rest views were used for investigating view synthesis
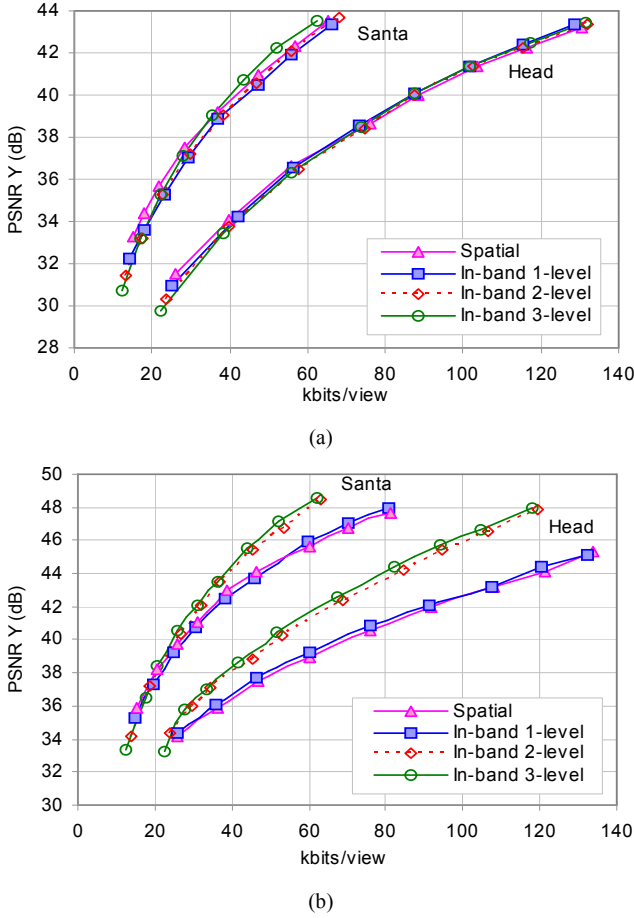
(a)

(b)

Fig. 9. Rate-distortion performance of (a) the half-resolution images and (b) the quarter-resolution images.
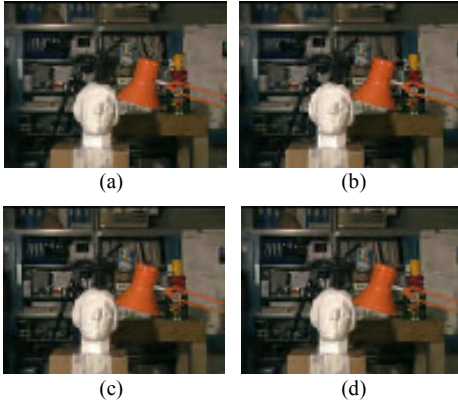

Fig. 10. The quarter-resolution reconstructed middle views of *Head* at 30kbits/view (a) block-based spatial prediction (PSNR Y=35.13). (b) DP spatial prediction (PSNR Y=34.87). (c) block-based 3-level in-band prediction (PSNR Y=35.94). (d) DP 3-level in-band prediction (PSNR Y=36.35).

efficiency. The DWT was implemented with biorthogonal 9/7 filter up to total 5 decomposition levels in luminance component and up to 4 decomposition levels in chrominance components. Block sizes of 8x8, 4x4 and 2x2 are chosen for in-band estimation in level 1, 2 and 3 respectively, while the

TABLE II
PERFORMANCE OF VARIOUS LEVELS OF THE IN-BAND PREDICTION

level	Head		Santa		Breakdancers	
	40	140	20	100	40	100
0*	31.04	40.32	33.25	42.63	33.82	41.71
1	30.96	40.37	33.22	42.34	33.37	41.13
2	30.88	40.44	33.13	41.86	33.38	41.16
3	30.80	40.85	33.08	42.53	33.49	41.61
4	30.45	40.52	32.89	42.12	33.47	41.54

* Decomposition level of 0 indicates the spatial prediction

size of 8x8 is used for spatial estimation. Estimation was performed using one the luminance component, and then the prediction information was used for chrominance components. Residual data was encoded with EBCOT coding. The chrominance components (U,V) were compressed separately but included within the desired target bits.

A. Decomposition Levels of In-band Prediction

In this section, the performance of in-band prediction at several decomposition levels is investigated with three multi-view images which contain the different natures and contexts. The Head sequence comprises many edges and occlusion areas, while the Santa sequence is a doll containing smooth edges. The geometry of these two sequences is parallel, but the geometry of the Breakdancers is non-parallel.

The luminance component was decomposed into 1-4 levels, and then the in-band prediction was applied. Subsequently the residuals of low-pass prediction results were further decomposed until 5 levels were achieved before applying EBCOT coding. The performances were compared among the prediction at various levels with the proposed DP approach. The results are tabulated in TABLE II.

At low bitrates, when the level of the decomposition for in-band prediction increases, the performance slightly decreases since most target bits are spent for coding side information, which includes the extra bits used for identifying the choice of critically-sampled subbands, instead of being used for coding the texture of the image. However, the performance of the in-band prediction at low bitrate can further be improved by including only the necessary disparity information into bitstream. At high bitrates, the performance increases as the level of the decomposition for in-band prediction increases. Interestingly, the extra bits used for identifying the choice of critically-sampled subbands affect the coding performance again when predicting at very coarse-resolution level as shown in the table that the quality of the 4-level decomposition slightly drops from the 3-level decomposition. This is because the number of critically-sampled subbands of the ODWT is proportional to $2^{2 \times level}$, so the number of these subbands is greatly increased from 64 subbands to 256 subbands when decomposing with level 3 and 4 respectively.

B. Compression Performance

Fig. 6 shows the comparison between predicting in spatial domain and in wavelet domain at 3-level decomposition. The results of the prediction in spatial domain and in wavelet domain are comparable. The in-band prediction improves the
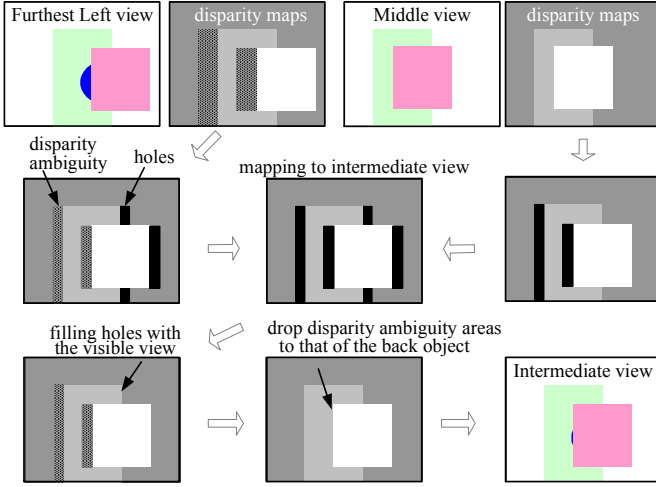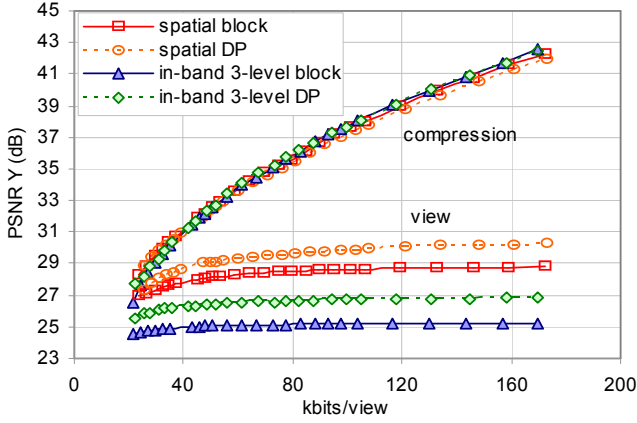
Fig. 11. The basic view synthesis approach.

Fig. 12. Rate-distortion performance of the compression and view synthesis for *Head* at full resolution.
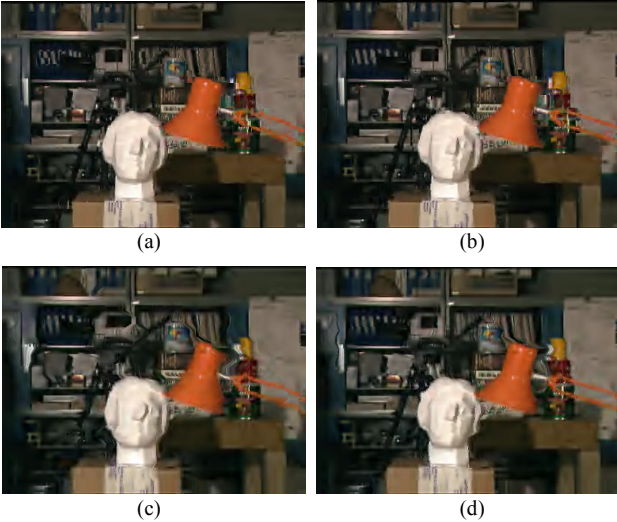

Fig. 13. The intermediate view of Head sequence between the left and middle views coded at 30 kbits/view. (a) block-based spatial prediction. (b) DP spatial prediction. (c) block-based 3-level in-band prediction. (d) DP 3-level in-band prediction.

performance at high bitrate, but gives a poorer quality at lower bitrates where more bits are used for coding side information,

including the choice of critically-sampled subbands. The examples of the reconstructed views are shown in Fig. 7. The performance of the block-based estimation is similar to that of the proposed DP. The prediction of high-pass subbands plays an important rule in increasing the compression efficiency.

The direct prediction in high-pass subbands gives a good edge result and no blocking artefacts [14]. Fig. 8 depicts the subjective results of Head sequence. At the same bitrate, the spatial prediction produces visually blocking artefacts, especially along the diagonal edges.

C. Spatial Scalability

The performance of the proposed scheme is compared to that of conventional spatial prediction by considering the image quality of the reduced-size images. The evaluation has been done by calculating the PSNR values of the reconstructed low-pass images at the same level of the decomposed original images. The average PSNR is displayed in Fig. 9 and the subjective results of the quarter size are shown in Fig. 10 Noticeably, the in-band prediction outperforms the spatial prediction but it does not improve much at very low bitrate as a result of side information. The in-band prediction at level 1 gains similar results of the quarter size as the spatial prediction, since the fractional values disparity vectors of low-pass subband are rounded to integer values. It is greatly affected at the object boundaries in which the discontinuity of depth visibly presents. Consequently, the optimal level for predicting subband-to-subband should be equal or higher (more decomposition) than the required resolution at the receiver. However, the quality of the reconstructed images at the full-resolution should not be deteriorated.

D. View Synthesis Performance

The view synthesis performance was investigated at full resolution and also coarser resolutions by taking off one of the reference views to be used for view synthesis assessment. The remaining views were then used for disparity/depth estimation by the traditional block-based estimation and the proposed DP. At the decoder, the disparity/depth is used with the decoded textures for synthesising the intermediate views. The disparity vectors from the low-pass subband prediction are utilized. They are linearly interpolated for the finer resolution and sub-sampled for the coarser resolution. The textures of the reference views are mapped to the intermediate views by disparity information. The textures of the regions attaining the disparity consistency are an average of textures from both nearest reference views, while the holes are filled with the texture of the visible view. However, the problem of filling the holes probably occur if one of the reference views is the furthest left or the furthest right view of multi-view sequence. Such areas might get the incorrect disparity, because no corresponding areas in other views. We synthesize the intermediate view in these positions by filling the tricky holes with the disparity values similar to the back objects as described in Fig. 11.

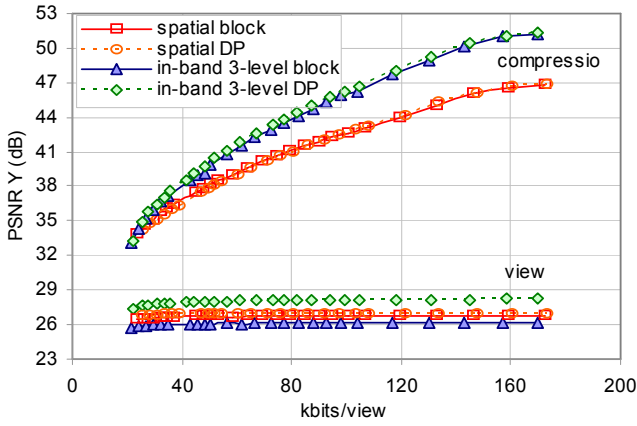*1) View Synthesis at Full Resolution:* The evaluation was

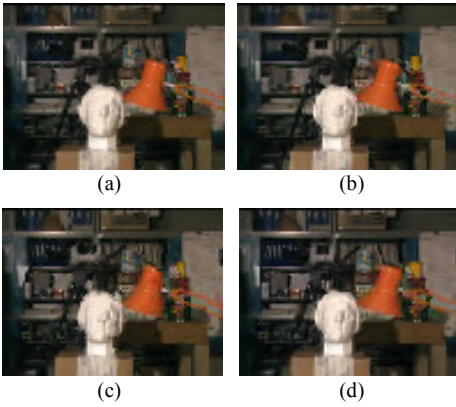Fig. 14. Rate-distortion performance of the compression and view synthesis for *Head* at quarter resolution.

Fig. 15. The intermediate view of Head between the left and middle views at quarter-resolution. (a) block-based spatial prediction. (b) DP spatial prediction. (c) block-based 3-level in-band prediction. (d) DP 3-level in-band prediction
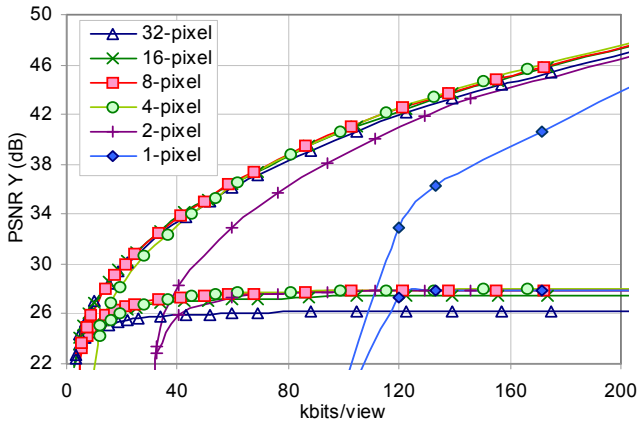

Fig. 16. Rate-distortion comparison of various resolutions of disparity map for *Head*.

made by comparing the virtual views to the original views. The average PSNR are shown in Fig. 12 and the examples of the subjective results are displayed in Fig. 13. Comparing between the in-band and spatial prediction, the spatial prediction outperform at full resolution, since the disparities are directly estimated at the full resolution. The in-band prediction

generates high mistakes due to the linear interpolation of the disparity/depth maps. The Head results synthesised from the in-band prediction show distorted curves at the depth discontinuity, e.g. the books located behind the head or the lamp.

Comparing between the block-based estimation and the DP, the disparity/depth maps estimated by the DP can significantly improve the view synthesis performance, especially at the areas where depth discontinuity is not present, e.g. the background, because the DP provides smooth disparity. For non-parallel sequence, the synthesized views of Breakdancers from the DP are significantly better than the results from the block-based estimation both in-band and spatial prediction.

2) View Synthesis at Coarse Resolution: In the limited display size, the view synthesis process exploited the small reconstructed image, and the disparity/depth maps are sup-sampled (for the case that the available disparity/depth map contains higher resolution) or are interpolated (for the case that the available disparity/depth map contains lower resolution) into the proper resolution. Fig. 14 and Fig. 15 show the objective and subjective results of view synthesis of the Head sequence at quarter resolution respectively. The average PSNR of the spatial prediction and 3-level in-band prediction by block-based matching are comparable, while the in-band prediction by proposed DP greatly improves the view synthesis at small image size. This is the result of the more accurate disparity estimation of the DP than that of the block-based algorithm. Note that the experimental results of the Santa and Breakdancers sequences show the similar trend of rate-distortion performance in which the DP improves the synthesis performance up to 3 dB.

E. Compression versus View Synthesis

The best algorithm for multi-view image coding is a trade-off between the required data size and quality of virtual view construction. Some disparity estimation algorithms support the compression requirement but are not suitable for view synthesis, e.g. block-based matching. Some enhanced approaches offer the correct depth results but require more data which compromises compression efficiency. In this paper, the hybrid in-band prediction and the DP have been proposed to improve the low-pass subband prediction. The proper filters are selected so as to increase the geometric coding performance, while the estimated disparities/depths present more precisely thereby achieving view synthesis.

The relationship between the compression and the view synthesis was investigated by considering the bit allocation for coding texture and disparity/depth information. Basically, the disparity/depth is compressed with lossless coding, while the textures are compressed with lossy coding. The bits allocated for disparity/depth coding can be adjusted by varying the resolution of the disparity/depth maps. Theoretically, the higher-resolution disparity/depth maps give the better quality of the synthesised results; therefore, the performance of the view synthesis can be improved by using the dense disparity/depth maps. However, the denser disparity/depth

map needs higher amount of bits for coding which may lead poor quality of textures for each particular bit target.

The simulation results are illustrated in Fig. 16 with various disparity/depth resolutions. It clearly proves that the higher-resolution disparity/depth maps generate better quality of the view synthesis. However, this benefit is achieved only in high bitrate because the dense disparity/depth maps require the large space for storing or transmitting. Interestingly, the highest compression performance achieved when using the 4- or 8-pixel resolution. The quality of the reconstructed views with the coarser-resolution disparity maps, e.g. 16- and 32-pixel resolution, decreases since it produces the poorer prediction which causes high energy at the high-pass views and deteriorates the efficiency of entropy coding. From the simulation of these three multi-view test sequences, it can be concluded that the 8-pixel resolution gives the best results for both compression and view synthesis.

From our experiments above, it can be concluded that the proposed hybrid prediction reduces less the energy of the geometric high-pass subband comparing to the traditional filters, Haar and 5/3 wavelet filter. It leads the efficiency of the disparity compensated view filtering. Next, the accurate disparity map by using DP approach is applied as a criterion for the hybrid prediction. Although, the PSNR values of the reconstructed images of some sequences may be slightly deteriorated (up to 0.25 dB), their subjective qualities are comparable and the viewers do not feel the difference. In contrast, the performance of the view synthesis is significantly improved (up to 1.5 dB).

The in-band prediction is applied to the geometry lifting scheme. The simulation results show the improvement of smaller scale of reconstructed images, especially at very small size (up to 4.5 dB). By truncating the high-pass subbands, the information for reconstructing the lower-resolution image from the in-band prediction does not lose as that from the spatial prediction. Therefore, this feature highly benefit the very small display. However, the disparity information employed for view synthesis is gained from only the low-pass subbands which cause the poor quality of full-resolution virtual views. Consequently, the disparity map from the DP approach is applied to the low-pass prediction instead of the block matching. The view synthesis is highly improved as illustrated both the subjective results and the objective results (up to 1.4 dB at full resolution and 1.8 dB at lower resolution). The main drawback of the proposed I-DCVF is that the CODWT process is required to generate the overcomplete high-pass subbands at the decoder. Consequently, the complexity of the codec is higher than the conventional MCTF.

VI. CONCLUSION

A novel in-band disparity compensation for the multi-view image coding based on wavelet lifting scheme is proposed in this paper. Hybrid prediction is exploited to deal with the occlusion problem. This prediction operates in wavelet domain efficiently providing resolution scalability. Additionally, the proposed I-DCVF can exploit different filters and estimation parameters for each resolution level. Perceived view synthesis quality is enhanced by using the DP for disparity estimation and filter selector in low-pass subband prediction. SNR scalability is provided by using EBCOT coding. As a result, the proposed codec is both scalable and efficient for multi-view image compression as well as achieve satisfactory view synthesis. Our experiments show comparable results at full resolution and significant improvement at coarser resolution.

REFERENCES

[1] M. E. Lukacs, "Predictive Coding of Multi-Viewpoint Image Sets," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 1986, pp. 521-524.

[2] J. Liu and R. Skerjanc, "Construction of Intermediate Pictures for a Multiview 3D System," in *Proc. SPIE Stereoscopic Displays and Applications III*, vol. 1669, 1992, pp.10-19.

[3] B. Pesquet-Popescu and V. Bottreau, "Three-dimensional Lifting Schemes for Motion Compensated Video Compression," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, May 2001, pp. 1793-1796.

[4] V. Bottreau, M. Benetiere, B. Felts, B. Pesquet-Popescu, "A Fully Scalable 3D Subband Video Codec," in *Proc. IEEE Int. Conf. Image Processing*, Vol 2 , Oct. 2001, pp. 1017 – 1020.

[5] L. Luo, J. Li, S. Li, Z. Zhuang and Y. Zhang, "Motion Compensated Lifting Wavelet and Its Application in Video Coding," in *Proc. IEEE Int. Conf. Multimedia and Expo*, Aug 2001, pp. 481-484.

[6] A. Secker and D. Taubman, "Lifting-Based Invertible Motion Adaptive Transform (LIMAT) Framework for Highly Scalable Video Compression," *IEEE Trans. Image Processing*. vol. 12, no. 12, Dec 2003, pp. 1530-1542.

[7] N. Anantrasirichai, C. Nishan Canagarajah, and David R. Bull, "Lifting-Based Multi-View Image Coding," in *Proc. IEEE Int. Sym. Circuits and Systems*, vol.3, Kobe Japan, May 2005, pp.2092-2095.

[8] N. Anantrasirichai, C. Nishan Canagarajah, David W. Redmill and David R. Bull, "Dynamic Programming for Multi-view Disparity/Depth Estimation," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol.2, 2006, pp. 269-272.

[9] R. Zaciu and F. Bellijemine, "A Compression Method for image sequences," in *Proc. IEEE Int. Conf. Consumer Electronics*, 1994, pp. 230-231.

[10] H. -W. Park and H.-S. Kim, "Motion Estimation Using Low-Band-Shift Method for Wavelet-Based Moving-Picture Coding," *IEEE Trans. Image Processing*, vol.9, no.4, April 2000, pp. 577-587.

[11] X. Li, L. Kerofski, and S. Lei, "All-phase motion compensated prediction in the wavelet domain for high performance video coding," in *Proc.IEEE Int. Conf. Image Process.*, vol. 3, Thessaloniki, 2001, pp. 538–541.

[12] C. Mayer and S. Albert, "Spatially Scalable Coding with In-band Prediction," in *Proc. SPIE Visual Communications and Image Processing*, vol. 5022, Santa Clara, USA, Jan 2003, pp.973-983.

[13] G. Van der Auwera, A. Munteanu, P. Schelkens and J. Cornelis, "Bottom-up motion compensated prediction in the wavelet domain for spatially-scalable video", *IEE Electronics Letters*, vol. 38, no. 21, Oct. 2002, pp. 1251-1253.

[14] J. Ye and M. Schaar, " Fully Scalable 3-D Overcomplete Wavelet Video Coding using Adaptive Motion Compensated Temporal Filtering," in *Proc. SPIE Visual Communications and Image Processing*, vol.5150, June 2003, pp.1169-1180.

[15] Ohta, Y., Kanade, T "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1985, pp. 139-154.

[16] W. Swelden, "The Lifting Scheme: A Construction of Second Generation Wavelets," *SIAM Journal of Mathematical Analysis*, vol. 29, no. 2, Mar 1998, pp. 511-546.

[17] W. Sweldens, "The lifting scheme: A new philosophy in biorthogonal wavelet constructions," *Wavelet Applicat. Signal Image Process.*, vol. 2569, 1995, pp. 68–79.

[18] F. A. M. L. Bruekers and A. W. M. van den Enden, "New networks for perfect inversion and perfect reconstruction," *IEEE J. Select. Areas Commun.*, vol. 10, 1992, pp. 129–137.

[19] J. Ohm, "Three dimensional subband coding with motion compensation," *IEEE Trans. Image Processing*, vol.3, Sep 1994, pp 559-571.

[20] Abhijeet V. Golwelkar and John W. Woods, "Scalable Video Compression using Longer Motion Compensated Temporal Filters," in *Proc. SPIE Visual Communications and Image Processing*, vol.5150, 2003, pp.1406-1416.

[21] W. Yand, F.Wu, Y. Lu, J.Cai, K.Ngan, and S. Li, "Scalable multi-view video coding using wavelet," in *Proc. IEEE Int. Sym. Circuits and Systems*, vol.3, Kobe Japan, May 2005, pp.6078-6081.

[22] B. Girod, C. -L. Chang, P. Ramanathan, and X. Zhu, "Light field compression using disparity-compensated lifting," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol.4, 2003, pp. 760-763.

[23] C. -L. Chang, X. Zhu, P. Ramanathan, and B. Girod, "Light field compression using disparity-compensated lifting and Shape Adaptation," *IEEE Trans Image Processing*, vol.15, no.4, April 2006, pp. 793-806.

[24] N. Anantrasirichai, C. Nishan Canagarajah, and David R. Bull, "Multi-View Image Coding with Wavelet Lifting and In-Band Disparity Compensation," in *Proc. IEEE Int. Image Processing*, vol. 3, Italy, Sep 2005, pp.33-36.

[25] Y. Andreopoulos, A. Munteanu, J. Barbarien, M. Schaar, J. Cornelis and P. Schelkens, "In-Band Motion Compensated Temporal Filtering," *Signal Processing: Image Communication*, vol.19, 2004, pp. 653 – 673.

[26] Andreopoulos, Y.; Munteanu, A.; Van der Auwera, G.; Cornelis, J.P.H.; Schelkens, P.; Complete-to-overcomplete discrete wavelet transforms: theory and applications," *IEEE Trans. Signal Processing*, VOL. 53, NO. 4, APRIL 2005,pp. 1398 – 1412.

[27] G. Piella and H.J.A.M. Heijmans, "Adaptive Lifting Schemes with Perfect Reconstruction," *IEEE Trans. Signal Processing*, vol. 50, no. 7, Jul 2002, pp.1620-1630.

[28] N. Mehrsereht, D. Taubman, "Adaptively Weighted Update Steps in Motion Compensated Lifting Based Scalable Video Compression", in *Proc. IEEE Int. Image Processing*, vol.3, Sept 2003, pp.771-774.

[29] Adams, D. Michael, (2002). The JPEG-2000 Still Image Compression Standard (Last Revised: 2002-12-25). ISO/IEC JTC 1/SC 29/WG 1N 2412.

[30] E. Chan, A.A. Rodriguez, R. Gandhi and S. Panchanathan, "Experiments on block-matching techniques for video coding," *Multimedia System*, Springer Berlin, 2005, pp. 228-241.

[31] K. Nalayama and S. Shimojo, "Da Vinci Stereopsis: Depth and Subjective Occluding Contours from Unpaired Image Points," *Vision Research*, vol. 30, 1990, pp. 1811-1825.

[32] I. Cox, "A maximum likelihood N-Camera stereo algorithm," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 733-739.

[33] I. Cox, S. Hingorani, and S. Rao, "A Maximum Likelihood Stereo Algorithm," Computer Vision and Image Understanding, Vol. 63, No. 3, May 1996, pp. 542–567.

[34] Peter N. Belhumeur, David Mumford, "A Bayesian Treatment of the Stereo Correspondence Problem Using Half-Occluded Regions," Proceedings of CVPR '92, 15-18 June 1992, pp. 506 – 512.

[35] N. Grammalidis and M. Strintzis, "Disparity and Occlusion Estimation in Multiocular Systems and Their Coding for the Communication of Multiview Image Sequences," IEEE Trans. on Circuits and System for Video Technology, Vol.8, No.3, Jun 1998, pp. 328-344.

[36] D. Scharstein, R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *Inter. Journal of computer vision*, Springer, vol.47, Nov 2002, pp. 7-42.

[37] Middlebury College.   Available: http://vision.middlebury.edu/~schar/stereo/web/results.php

N. Anantrasirichai (S'04-M'07) received the B.E. (Hons.) degree in Electrical Engineering from Chulalongkorn University, Thailand, the M.E. degree in Telecommunications from Asian Institute of Technology, Thailand and the Ph.D. degree in Electrical & Electronic Engineering, University of Bristol, U.K. She is currently a research assistant at the Centre of Communications Research at the University of Bristol. Her research interests include image and video coding, multi-view processing and distributed video coding.

C. Nishan Canagarajah (M'95) received the B.A. (Hons.) and Ph.D. degrees in DSP techniques for speech enhancement, both from the University of Cambridge, Cambridge, U.K. He is currently a Professor of Multimedia Signal Processing at the University of Bristol, U.K. He was previously an Research Assistant and Lecturer at Bristol, investigating DSP aspects of mobile radio receivers. His research interests include image and video coding, image segmentation, content based video retrieval, 3–D video and image fusion. He is widely supported in these areas by industry, EU and the EPSRC. He has been involved in a number of EU FP5 and FP6 projects where the team has been developing novel image/video processing algorithms. He has published more than 160 papers and two books. Prof. Canagarajah is a member of the EPSRC Peer Review College.

David W. Redmill received the B.A. (Hons.) degree in electrical and informational sciences in 1991 and the Ph.D. degree in 1995, both from the department of Engineering, University of Cambridge, UK. He is currently a lecturer at University of Bristol. His research interests include low bit-rate image and video coding for noisy channel, as well as multi-view image/video processing.

David R. Bull (M'94–SM'07) holds the Chair in Signal Processing at the University of Bristol. He leads the Signal Processing activities in the Centre for Communications Research and is Director of the Bristol Vision Institute. He is also co-founder and Chairman of ProVision Communication Technologies Ltd. Prior to his current appointments he was a Systems Engineer at Rolls Royce and subsequently a Lecturer at Cardiff University. He has published approximately 400 papers and two books and has also acted as an independent consultant to numerous international organisations in the fields of video coding and signal analysis. Prof. Bull is a Fellow of the IEE and a Chartered Engineer, U.K.