Overlapped Block Disparity Compensation with Adaptive Windows for Stereo Image Coding

Woontack Woo and Antonio Ortega

Abstract—**In this letter, we propose a modified overlapped block-matching (OBM) scheme for stereo image coding. OBM has been used in video coding but, to the best of our knowledge, it has not been applied to stereo image coding to date. In video coding, OBM has proven useful in reducing blocking artifacts (since multiple vectors can be used for each block), while also maintaining most of the advantages of fixed-size block matching. There are two main novelties in this work. First, we show that OBM techniques can be successfully applied to stereo image coding. Second, we take advantage of the smoothness properties typically found in disparity fields to further improve the performance of OBM in this particular application. Specifically, we note that practical OBM approaches use noniterative estimation techniques, which produce lower quality estimates than iterative methods. By introducing smoothness constraints into the noniterative DV computation, we improve the quality of the estimated disparity as compared to standard noniterative OBM approaches. In addition, we propose a disparity estimation/compensation approach using adaptive windows with variable shapes, which results in a reduction in complexity. We provide experimental results that show that our proposed hybrid OBM scheme achieves a PSNR gain (about 1.5–2 dB) as compared to a simple block-based scheme, with some slight PSNR gains (about 0.2–0.5 dB) in a reduced complexity, as compared to an approach based on standard OBM with half-pixel accuracy.**

Index Terms—**Disparity compensation, half-pixel search, MRF model, overlapped block matching, stereo image.**

I. INTRODUCTION

The overall bit rate of a coded stereo image pair can be reduced by exploiting the inherent similarity between the two images. As in video coding, a popular approach to exploit this similarity is to find a disparity vector (DV) field for one of the images and then transmit the disparity compensated difference (DCD). In addition, as in video, the DV field is normally computed for fixed-size blocks rather than pixels or features because fixed-size block matching (FSBM) achieves reasonable performance while being simple to implement [1], [2]. In a FSBM scheme the encoder transmits, as overhead, the vectors for each block, along with the corresponding DCD. The DCD energy is lower for smaller blocks, at the cost of higher overhead, and thus, the block size to be used for matching can be selected by analyzing the corresponding rate-distortion (RD) tradeoffs [3].

However, FSBM has several well-known drawbacks. First, inaccurate disparity estimation and compensation (DE/DC) can result because: 1) a single disparity is transmitted for each block; 2) the accuracy of the DV (e.g., integer-pixel resolution) may not be sufficient to describe the actual disparity; and 3) intensity noise effects and occlusions lead to disparity vectors that minimize the DCD energy but do not correspond to the true disparity. Note that DV field inaccuracies due to items 1) and 2) can result in higher DCD frame energy, and therefore, higher required transmission rate. In addition, in situations where it is desirable to estimate intermediate images from the stereo pair, a DV estimation that does not correspond to the true disparity [e.g., because of 3)] will lead to artifacts appearing in the intermediate view. Finally, FSBM tends to produce artifacts at the block boundaries in the predicted image, especially when low rate coding is used.

Therefore, efficient alternatives to FSBM for DE/DC have been a main focus of the research on stereo image/video coding since the pioneering work by Lukacs [2], [4]–[8]. However, the various proposed schemes have some shortcomings. For example, DE with a Markov random field (MRF) model can overcome the inconsistency of the DV field by taking advantage of disparity information of neighboring blocks [5], [9], [10]. Subspace projection is another way of estimating a smooth DV field [8]. However, both schemes have limitations because the smoothness constraint tends to increase the energy level of the DCD frame, and therefore, more bits may be required for coding. Meanwhile, the energy level of the DCD frame can be reduced using noninteger (half or quarter) pixel-based search at the cost of an increase in the rate of the DV field. The remaining problem, the blocking artifacts, can be reduced by adopting various other methods, such as post-processing, segmentation-based estimation/compensation, etc. However, many post-processing algorithms degrade the quality of the whole image, as well as the block boundaries. While segmentation-based methods may be appealing, they come at the cost of the increase in overhead to describe the structure of the segmentation [6], [11], [12]. Moreover, algorithms for robust automatic segmentation may not be available.

As an alternative, we propose overlapped block disparity compensation (OBDC), an extension of the overlapped block motion compensation (OBMC) for video coding, where OBMC reduces blocking artifacts by linearly combining the predictions generated using multiple motion vectors, including a block's motion vector as well as its neighbors [13]–[16]. The main drawback of OBMC techniques is that optimized motion field

---

Manuscript received March 15, 1999; revised September 30, 1999. This work was supported in part by the National Science Foundation under Grant MIP-9502227 (CAREER), by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, by the Annenberg Center for Communication, University of Southern California, and by the California Trade and Commerce Agency. This paper was recommended by Guest Editor K. N. Ngan.

W. Woo was with the Integrated Media Systems Center, Department of Electrical Engineering Systems, University of Southern California, Los Angeles, CA 90089 USA. He is now with ATR MIC Laboratories, Kyoto 619-0288, Japan (e-mail: wwoo@mic.atr.co.jp).

A. Ortega is with the Integrated Media Systems Center, Department of Electrical Engineering Systems, University of Southern California, Los Angeles, CA 90089 USA (e-mail: ortega@sipi.usc.edu).

Publisher Item Identifier S 1051-8215(00)02013-9.

estimation requires relatively complex iterative schemes to resolve the noncausal spatial interaction among motion vectors of neighboring blocks. For practical implementations, in order to reduce this estimation complexity, "two-step procedures" can be adopted, i.e., an initial motion vector field is estimated using FSBM and then the motion field is refined to improve the encoding performance. For example, block matching (BM) [13] or windowed BM [17] can be first applied for motion estimation, without considering the effects of neighboring blocks, and then OBMC is performed only for motion compensation. However, motion fields obtained with these techniques can sometimes be far from the optimal achievable performance since the motion vectors are computed to minimize the energy of a residue based on nonoverlapped compensation, while overlapped compensation will be used to compute the actual residue. As an alternative, modified noniterative OBMC schemes, such as raster scan OBMC [18] or checkerboard scan OBMC [19], have been proposed to achieve reasonable coding performance, without requiring a complex iterative estimation. The basic idea in these schemes is to perform an overlapped estimation based on those blocks that are already known (according to some predetermined scan). In this way, at least some of the neighboring blocks (those already scanned) can be taken into account in the motion estimation.

In this work, we show how the performance of an OBDC scheme can be improved by introducing smoothness constraints into the disparity estimation. These smoothness constraints allow us to reduce the cost of transmitting the disparity field (since it will be more correlated and we use differential coding techniques). Moreover, we will see how a smooth DV field results in reduced residue frame energy even when using a simple noniterative search technique. Specifically, we will show how the performance of simple two-step approaches [17] can be improved by introducing MRF constraints in the estimation process. We also show that it is beneficial to use variable window shapes within the OBMC framework. A fixed shape of the window as used in OBMC ignores the fact that OBMC usually works when the estimation is unreliable. In addition, while spreading compensation errors tends to reduce blocking artifacts, it might degrade compensation efficiency, particularly for those blocks that can be compensated effectively without OBMC. Based on these observations, OBMC can be turned off in cases where a reliable estimation exists. As an example, assume a uniform (foreground) block is surrounded by background blocks. Then, since the motion vectors of foreground and background will be quite different, it is doubtful that OBMC can improve performance over plain motion compensation. In our approach, we rely on the MRF information to determine the reliability of the estimate and use this information to determine which, if any, of the neighboring blocks should be used to perform an overlapped disparity compensation.

In this letter, we show that OBM is also useful for stereo image coding, and we demonstrate that it clearly outperforms standard block-matching techniques. The estimation of the DV field is done based on a noniterative approach. In our proposed OBDC, the overall encoding gain is improved with respect to a direct implementation of OBM using two techniques. First, we introduce a *DE with a modified MRF model* that results in a smooth DV field, without excessively increasing the energy level of the prediction error, while reducing the bit rate for the DV field. Then, given a smooth DV field, we propose a *selective OBDC* that reduces blocking artifacts and energy level of the DCD frame by adaptively changing the shape of the OBM window to prevent the oversmoothing problem.

To verify the effectiveness of our approach, we compare the RD performance of the proposed OBDC scheme with various FSBM-based DE/DC schemes, such as simple FSBM, FSBM with MRF model, and OBDC (direct extension of OBMC). According to our experimental results, the proposed scheme achieves about 1.5–2-dB gain compared to FSBM, about 1–1.5-dB gain compared to OBDC (based on conventional OBMC), and about 0.2–0.5 dB gain compared to OBDC with half-pixel accuracy.

This letter is organized as follows. In Section II, we describe the proposed two-step hybrid scheme, the modified OBM with MRF model and half-pixel search. In Section III, we provide some experimental results to compare the effectiveness of the proposed scheme.

II. MODIFIED OVERLAPPED BLOCK MATCHING

A. Definitions and Notation

We denote F_1 and F_2, the reference and target images, respectively. Our goal is to predict F_2 based on both F_1 and the corresponding DV field. Then, we encode both the DV field and the prediction error, instead of F_2. Let the target image F_2 be segmented into nonoverlapping square blocks, i.e., $F_2 = \{f_{ij}^2, (i, j) \in \Omega\}$, where f_{ij}^2 denotes the (i, j)th block and Ω represents a discrete and rectangular lattice, i.e., $\Omega = \{(i, j) | 0 \leq i < N_x, 0 \leq j < N_y\}$. N_x and N_y are the number of blocks in vertical and horizontal directions, respectively. Each block has a DV, v_{ij}, and thus the resulting blockwise DV field can be represented as $V = \{v_{ij}, (i, j) \in \Omega\}$. Thus, the estimate corresponds to a block in the reference image along v_{ij}, i.e., $\hat{f}_{ij}^2 = f_{ij \ominus v_{ij}}^1$, where $\oplus$ denotes a blockwise estimation. Note that $f_{ij \ominus v_{ij}}^1$ will no longer necessarily be on the grid Ω since arbitrary integer-pixel (or half-pixel) displacements are possible.

We define the (i, j)th enlarged block in the target image as s_{ij}^2 so that the windowed sum over the image is identical to the original image, i.e., $F_2 = \{W \cdot s_{ij}^2, (i, j) \in \Omega\}$, where W denotes a matrix corresponding to the enlarged window function. The usual width of the enlarged window is twice that of the original block, e.g., f_{ij} is a square block of $B \times B$, while W has a size of $2B \times 2B$, where B denotes the width (height) of the block. Note that conventional FSBM can be considered as an OBM with a rectangular function, e.g., s_{ij}^2 equals f_{ij}^2, and thus, each target block is estimated from only one block in the reference image along the disparity vector.

In OBM-based schemes, each target block is estimated and compensated as a windowed-sum of a block and its neighboring blocks along the corresponding DV's. In our experiments, we adopt the bilinear window, as shown in Fig. 1. In general, as shown in Fig. 1(a), the overlapped window W is designed to decay toward the boundaries on the assumption that blockwise

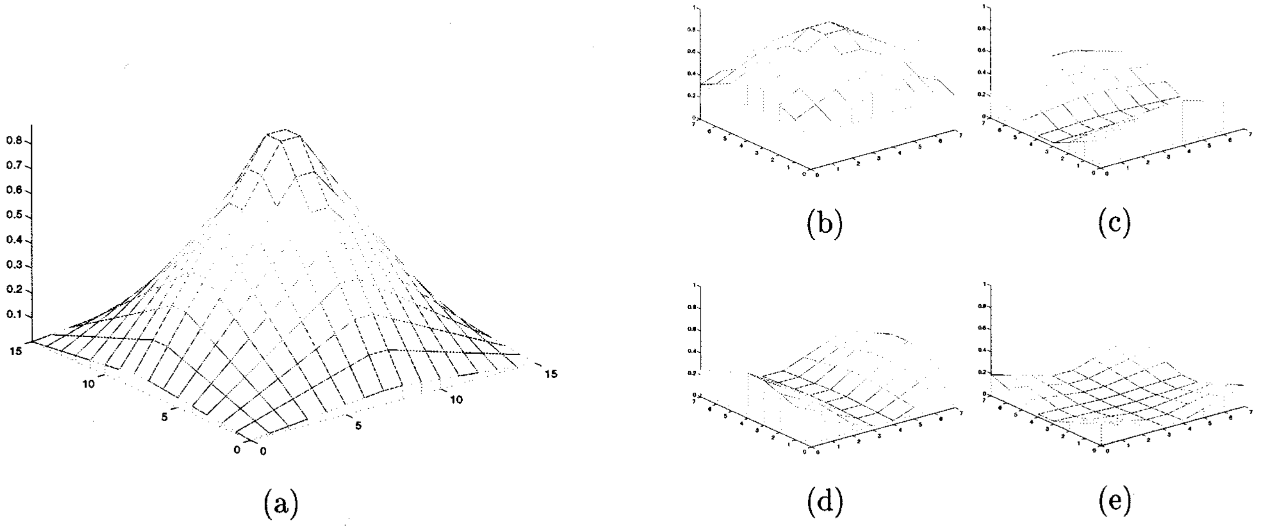


Fig. 1. Bilinear window function for the overlapped block matching and its combined weighting matrices. It shows how much the directional components of overlapping windows affect the compensation of a block (8 × 8). (a) Bilinear OBM window (16 × 16). (b) Main window ($\hat{f}^2_{ij}$). (c) Horizontal direction window for ($\hat{f}^2_{ij(\mathrm{N})}$, $\hat{f}^2_{ij(\mathrm{S})}$). (d) Vertical direction window for ($\hat{f}^2_{ij(\mathrm{W})}$, $\hat{f}^2_{ij(\mathrm{E})}$). (e) Corner area window for ($\hat{f}^2_{ij(\mathrm{NW})}$, $\hat{f}^2_{ij(\mathrm{NE})}$, $\hat{f}^2_{ij(\mathrm{SW})}$, $\hat{f}^2_{ij(\mathrm{SE})}$). The subscripts (N, W, S, E) denote locations of quadrants of a block, i.e., north, west, south, and east, respectively. In the proposed scheme, each portion of the window is added or removed according to ϕ_{ij}, the reliability of a neighboring DV.

estimation error increases as a pixel moves away from the block center, and the increase is symmetric with respect to the block center [20]. Typical selections for the overlapped windows are the sinusoidal and the bilinear windows. An optimal shape for the overlapped window can also be considered, but the resulting improvement is not a significant one, given the proportional increase in the computational complexity [14]. Fig. 1(b)–(e) show how much the directional components of overlapping windows affect the compensation of a block. The selected separable bilinear window $W_{m,n} = W_m \times W_n$, where W_m and W_n denote the separable vertical and horizontal windows, respectively, i.e.,

$$W_m = W_n = \begin{cases} \dfrac{m+0.5}{B}, & 0 \le m < B \\ W_{2B-m-1}, & B \le m < 2B. \end{cases} \quad (1)$$

Let ϕ^0_{ij} and ϕ_{ij}, respectively, denote the initial and final class of the ijth block. Here, $\phi_{ij} = \{0, 1, 2\}$ denote, respectively, block disparity compensation, OBDC, and non-DC (i.e., intra-coded). In the proposed scheme, each portion of the window is added or removed according to the reliability of a neighboring DV, ϕ_{ij}.

B. Disparity Estimation Using Overlapped Windows

Fig. 2 shows the DE using an overlapped window and a smoothness constraint. In general, the DV field obtained by the overlapped window is not likely to be smooth because the MSE/MAE-based prediction is sensitive to various noise effects such as intensity variation. For example, the two images in a pair may have slightly different intensity levels due to the camera noise and lighting condition. In addition, the lack of texture and/or repetitive texture may prevent consistent estimation.

Therefore, we first estimate a smooth DV field without iteration by introducing a causal MRF model. In general, conventional MRF-based schemes have high computational complexity because they require several (stochastic) iterations to estimate an optimal (pixelwise) dense DV field [21]. Thus, we propose a simplified blockwise DE scheme, which estimates a smooth DV field without complicated iterations by considering only blocks in a first-order causal neighborhood. The resulting smooth DV tends to reduce the bit rate for the DV field because the same neighborhood is used for encoding the DV field, i.e., the difference between v_{ij} and the median of its causal neighborhood is encoded using DPCM. As a result, estimating a smooth DV field contributes to reducing side information, which is especially essential at low rate coding.

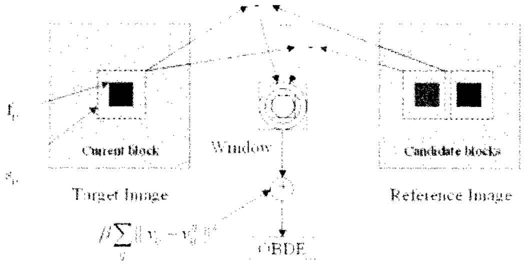


Fig. 2. Disparity estimation based on block matching with an enlarged window. In the target image, shaded and dashed areas correspond to a block, f_{ij}, and an enlarged block, s_{ij}, respectively.

We use the formulation in [5], [22] for the MRF-based DE. For the OBM, we set the weighting for the reliability measure, ϕ_{ij}, equal to zero and separately determine ϕ_{ij} according to the prediction error level. For example, a block with higher prediction error than a threshold is selected as an OBDC candidate, i.e., $\phi^0_{ij} = 1$, if $|\mathrm{DCD}| > T_\phi$, where T_ϕ denotes a threshold value. The DCD is calculated by taking the difference between the predicted block and the original block in the target image, i.e., $\mathrm{DCD} = f^2_{ij} - \hat{f}^2_{ij}$. Then, for the block with $\phi^0_{ij} = 1$, we set $\phi_{ij} = 1$, if the energy level of the DCD using OBDC is lower as compared to block-based (nonoverlapped) DC. However, for a block having the same DV as its neighbors, we set $\phi_{ij} = 0$, since there is no RD gain by OBDC when the disparity vectors of the neighboring blocks are the same. Meanwhile, if there is no gain from block DC or OBDC, we set $\phi = 2$ and then encode the original intensity block.

For DE based on overlapping block, the corresponding cost function in [5] has to be changed as follows:

$$\begin{aligned} U(V|F_1, F_2) &= U(F_2|F_1, V) + U(V) \\ &\propto \sum_{(i,j)\in\Omega} \Bigg\{ \|(1-\alpha)W \times (s_{ij}^2 - s_{ij\ominus v_{ij}}^1)\|^2 \\ &\quad + \alpha \sum_{\eta} \|v_{ij} - v_{ij}^{\eta}\|^2 \Bigg\} \end{aligned} \tag{2}$$

where η denotes a neighborhood and $\oplus$ represents a blockwise compensation. The weighting constant $(0 \leq \alpha \leq 1)$ controls the degree of smoothness. Each term of the right side in (2) represents the constraint of the similarity between stereo pair for a given disparity and an *a priori* assumption on the smoothness of the DV field. Note that settings $\alpha = 0, W = I$, and $s_{ij} = f_{ij}$ in (2) correspond to conventional FSBM, which only assumes that the image intensities in the stereo pair are similar along the DV.

Also note that, in the proposed scheme, the choice of model parameters is relatively robust. For example, a small fixed weight (e.g., $\alpha = 0.1$) is sufficient for the smoothness term for most images, though an optimal value of α can be selected by Lagrangian optimization, because the smoothness constraint is exploited only to avoid various local minima in DE. However, conventional MRF schemes, mainly employed in computer vision, require a more careful selection of an "optimal" set of weighting parameters, in order to provide good results.

To further reduce the prediction error, the DE/DC is performed in half-pixel accuracy, since the actual correspondence between the two images may not be aligned with integer-pixel location. Therefore, estimating/compensating the target image on the interpolated reference image along the disparity vectors helps to estimate a more accurate DV field and thus reduces the energy of the DCD frame.

C. Encoding with Selective OBDC

We selectively apply OBDC to reduce computational complexity and to increase compensation efficiency. This approach works because OBDC is efficient only when the energy level of the DCD block is significantly different from that of its neighboring blocks or when high frequency components exist in the DCD block. Note that we assume that the block can be compensated effectively without OBDC, if we set $\phi_{ij}^0 = 0$ for a given block. Thus, for $\phi_{ij} = 0$, we use a flat window to compensate the block. Then, OBDC is only considered for the block with $\phi_{ij}^0 = 1$. If $\phi_{ij} = 1$, we partially use each window to compensate the block, according to the reliability of the corresponding vector. Otherwise, i.e., if $\phi_{ij} = 2$, we encode the block in intramode, instead of the DCD block.

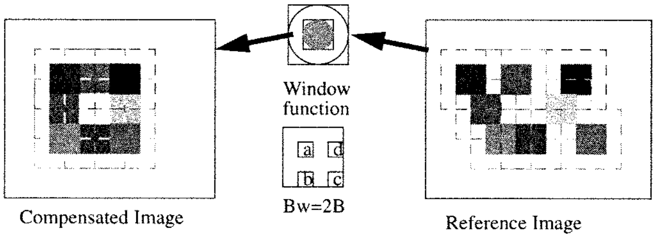


Fig. 3. Disparity compensation based on the overlapped block matching. Connected and dashed areas correspond to a block f_{ij} and an enlarged block s_{ij}, respectively.

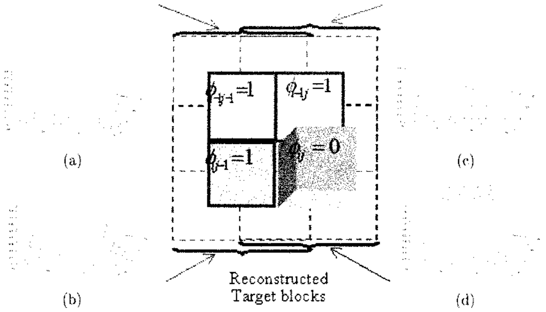


Fig. 4. Adaptive windowing for selective overlapped block disparity compensation. Given $\phi_{ij} = 0$ and $\phi_{i-1j-1} = \phi_{i-1j} = \phi_{ij-1} = 1$, OBM windows are changed adaptively according to the ϕ's. OBM windows for: (a) s_{i-1j-1}, (b) s_{i-1j}, (c) s_{ij-1}, and (d) s_{ij}.

As shown in Fig. 3, in OBDC, a target block is influenced by the nine overlapped blocks in the reference image along the corresponding disparity vectors. Thus, in the conventional approach, the compensation is performed by summing up the window-operated nine blocks.[1]

Note, however, that in the proposed scheme, OBDC is selectively applied to those blocks yielding higher prediction errors, i.e., $\phi_{ij} = 1$, while block DC is applied to all the others except $\phi_{ij} = 2$. If $\phi_{ij} = 0$, the neighboring blocks are regarded as having the same disparity vectors and thus the effects of the neighboring blocks are ignored. In addition, the adaptively changing window shape prevents the oversmoothing effects.

Fig. 4 shows an example of resulting window shapes, when $\phi_{ij} = 0$ and ϕ's of other blocks are one. Another example is shown in Fig. 5. Note that the block with $\phi_{ij} = 0$ affects neighboring blocks with $\phi = 1$. The nonsymmetrical approach helps prevent oversmoothing, while reducing blocking artifacts. Note also that the proposed scheme corresponds to the conventional OBDC, if all blocks are set to $\phi = 1$.

The encoding procedure based on the proposed selective OBDC is as follows

Step 0: The reference image is independently encoded using JPEG.

Step 1: The disparity is estimated using an enlarged bilinear window with $B_w = 2B$. The window function W is operated on the disparity-predicted difference, without considering DE errors of neighboring blocks, i.e., $\|W \times \{s_{ij}^2 - \hat{s}_{ij}^2\}\|$. The corresponding

---

[1] In the case where the window width and height are double those of the block, i.e., $B_w = 2 \times B$, one quarter of a block only depends on three neighboring blocks and itself. For example, each pixel in the upper left part (NW) of the target block $f_{ij(NW)}^2$ is compensated by the weighted sum of only four neighboring blocks.

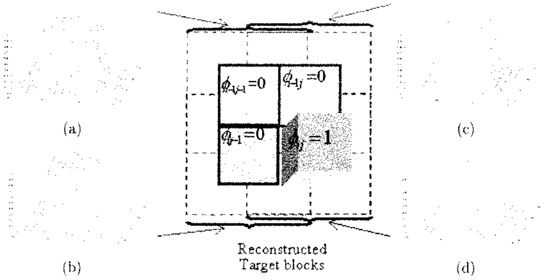


Fig. 5. Adaptive windowing for selective overlapped block disparity compensation. Given $\phi_{ij} = 1$ and $\phi_{i-1j-1} = \phi_{i-1j} = \phi_{ij-1} = 0$, OBM windows are changed adaptively according to the ϕ's. OBM windows for: (a) s_{i-1j-1}; (b) s_{i-1j}; (c) s_{ij-1}; and (d) s_{ij}.

DE cost is defined by adding a smoothness constraint as shown in (2). The estimation is performed in half-pixel accuracy.

Step 2: Given a DV, a block is determined as an OBDC candidate, i.e., $\phi^0_{ij} = 1$, if the energy level of the difference is larger than the threshold, i.e., $|\text{DCD}| = |f^2_{ij} - \hat{f}^2_{ij}| > T_\phi$. Otherwise, $\phi^0_{ij} = 0$.

Step 3: After DE with windowed BM, for a block with $\phi^0_{ij} = 1$, we compare the energy levels of block DC and OBDC. We select the block as an OBDC block, i.e., $\phi_{ij} = 1$, if the resulting DCD by OBDC has less energy than that of block DC. If there is no gain from either block DC or OBDC, i.e., $\phi_{ij} = 2$, the block is coded in intramode.

Step 4: OBDC is selectively performed for those blocks with $\phi_{ij} = 1$ by summing up all windowed compensation blocks.

Step 5: The resulting DV field and DCD frame are encoded using DPCM and JPEG, respectively. For DPCM of the DV field, we use a median predictor based on a predefined causal neighborhood.

At the decoder, the reference image is decoded first and then the target image is reconstructed according to ϕ_{ij}, e.g., by performing OBDC, block DC and non-DC. The final target image is reconstructed by adding information from the DCD for those blocks that have been predicted.

III. EXPERIMENTAL RESULTS

In this experiment, the right image is selected as a reference image and then a constant quantization factor ($Q_1 = 75$) is assigned for that reference image. Then, we measure the performance in terms of bit rates and peak signal-to-noise ratio (PSNR) of the encoded target image. Exhaustive search is performed within a search range of [0, ± 15] pixels in half-pixel accuracy. For the images that do not satisfy the parallax constraint, we search ± 2 pixels in vertical direction. In order to test the effectiveness of the proposed algorithm, we have simulated its performance for two pairs of stereo images; a synthesized scene, *Room,* and a natural scene, *Aqua.* The image sizes of the pairs are 256 × 256 and 288 × 360, respectively. In the following experiment, we choose a block size of 8 × 8 as a compromise between the overhead and the energy of the estimation error, and an enlarged window size of 16 × 16. The DV field is encoded using DPCM, i.e., the difference between v_{ij} and the median of its causal neighborhood is encoded using DPCM.

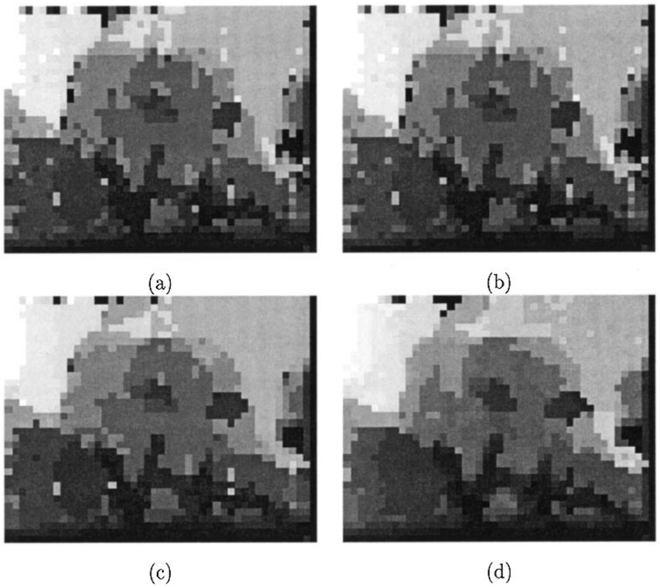


Fig. 6. DV fields for various disparity estimation methods. (a) FSBM (0.061 bpp). (b) DE with MRF (in [5]) (0.058 bpp). (c) DE with OBM (0.058 bpp). (d) The proposed scheme: OBD with MRF and half-pixel search (0.065 bpp). The combined method provides the most smooth and consistent DV field. Note that the DV field is linearly mapped into intensity between 0–255, and thus the brighter, the larger disparity value is.

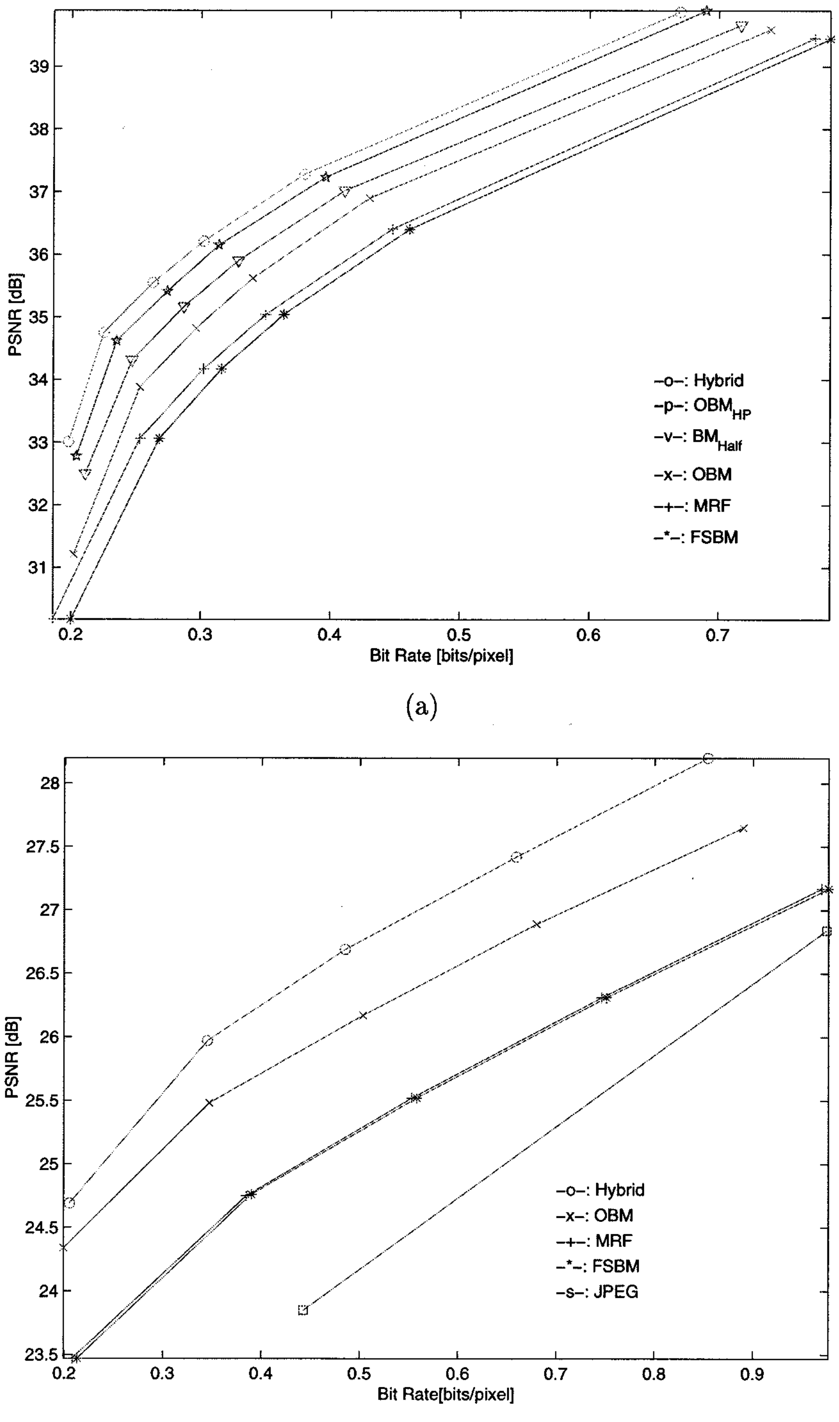
Fig. 6 compares the DV fields for various disparity estimation methods.[2] As expected, the MRF-based DE estimates a smoother DV field by the tradeoff between the spatial correlation in a stereo pair and the smoothness in the DV field. The proposed hybrid method provides the smoothest DV field, which appears to be more consistent with the actual DV field.

In our experiments, we use bilinear interpolation as a compromise to obtain the half-pixel precision intensity value, as used in most video coding standards. Clearly, the higher the subpixel accuracy (i.e., the larger displacement space), the greater the probability of finding a good match. Obviously, we cannot choose an arbitrarily small value because, as the subpixel accuracy increases, both the rate for the resulting DV field and the number of candidate blocks being compared in the search area increases.

Fig. 7(a) and (b) compare the corresponding RD plots for the two stereo pairs, *Room* and *Aqua,* respectively. The proposed selective OBDC scheme is compared to FSBM, MRF-based BM, BM with half-pixel accuracy, OBDC (based on OBMC), and OBDC with half-pixel accuracy in terms of PSNR and bit rate of the target image. Note that we provide a comparison only with disparity compensation schemes that are based on fixed block size, i.e., each of the proposed approaches would require transmission of a single disparity vector for each 8 × 8 block. Thus, comparisons with approaches allowing variable block sizes or region based compensation fall outside of the scope of this correspondence. Note also that for the natural image pair (*Aqua*), the RD gain of FSBM is relatively small, compared to that for *Room.*

---

[2] Stereo pairs and comparisons of DE results [Online]. Available: http://escalus.usc.edu/~wwoo/Stereo/



(a)

(b)

Fig. 7. The resulting RD plots (a) *Room* and (b) *Aqua*. We fix the quality of the reference images and then measure the performance in terms of bit rates and peak signal to noise ratio (PSNR) of the encoded target image (block size of 8 × 8, quality factor for the reference image $Q_1 = 75$): The proposed hybrid scheme is compared with FSBM, FSBM with MRF, FSBM with half-pixel accuracy, OBM and OBM with half-pixel accuracy. In the plot, "-p-," "-v-," and "-s-" denote the star, triangular, and square-mark lines, respectively.

As expected, the proposed selective OBDC results in an improved overall encoding performance, i.e., a lower bit rate for the DV field and DCD frame while maintaining a PSNR gain. In the proposed OBDC, the modified MRF model-based DE allows estimating a smooth DV field, while maintaining the energy level of the DCD frame (or slightly increases according to α). Then, the selective OBDC in half-pixel accuracy results in better compensation by reducing the energy level of the DCD frame as well as the blocking artifacts. Note that the main PSNR gain of the proposed scheme comes from properly combining various schemes. For example, adaptively changing the window helps reduce the energy level of the disparity compensation error when the disparity estimation is inaccurate. The smoothness term also reduces the increase in rate of the DV field due to the half-pixel search. The adaptively changing OBDC window reduces the oversmoothing problem, as well as blocking artifacts. As a result, the proposed selective OBDC scheme provides a higher PSNR, as well as better perceptual quality over other FSBM schemes such as MRF or OBMC, encoded at the same rate. According to our experimental results, the proposed modified OBDC scheme obtains a higher PSNR, about 1.5–2 dB as compared to block DC, about 1–1.5 dB over conventional OBMC schemes, and about 0.2–0.5 dB as compared to an OBMC with half-pixel accuracy.

The results of the proposed selective OBDC schemes can also be applied to video coding without loss of generality. It is worth noting that obtaining a smooth disparity is useful for multiview video coding since the robustness against noise can help reduce the temporal redundancy between two consecutive disparity fields. Note also that the resulting smooth DV field also helps generate intermediate-views with lower visual artifacts in the decoder. However, there remain several problems to be resolved. The performance can be increased further by adopting more elegant interpolation methods, such as an approximated ideal filter or Wiener filter [23], [24]. The overall encoding performance also could be improved by combining the proposed DE scheme and the dependent bit-allocation scheme proposed in [25], [26].

REFERENCES

[1] J. R. Jain and A.K. Jain, “Displacement measurement and its application in interframe image coding,” *IEEE Trans. Commun.*, vol. 29, pp. 1799–1808, Dec. 1981.

[2] M. E. Lukacs, “Predictive coding of multiviewpoint image sets,” in *Proc. ICASSP*, Oct. 1986, pp. 521–524.

[3] J. Ribas-Corbera and D. Neuhoff, “On the optimal block size for block-based motion-compensated video coders,” in *Proc. SPIE of VCIP*, vol. 3024, Feb. 1997, pp. 1132–1143.

[4] M. G. Perkins, “Data compression of stereopairs,” *IEEE Trans. Commun.*, vol. 40, pp. 684–696, Apr. 1992.

[5] W. Woo and A. Ortega, “Stereo image compression based on the disparity compensation using the MRF model,” in *Proc. SPIE VCIP*, vol. 2727, Mar. 1996, pp. 28–41.

[6] ——, “Stereo image compression based on the disparity field segmentation,,” in *Proc. SPIE EI-VCIP*, vol. 3024, Feb. 1997, pp. 391–402.

[7] M. B. Slima, J. Konrad, and A. Barwicz, “Improvement of stereo disparity estimation through balanced filtering: The sliding-block approach,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 913–920, Dec. 1997.

[8] H. Aydinoglu and M. H. Hayes, “Stereo image coding: A projection approach,” *IEEE Trans. Image Processing*, vol. 8, pp. 506–516, Apr. 1998.

[9] H. Jeong, W. Woo, C. Kim, and J. Kim, “A unification theory for early vision,” in *Proc. First Korea–Japan Joint Conf. on the Computer*, Oct. 1991, pp. 298–309.

[10] K. Illgner and F. Muller, “Motion estimation using overlapped block motion compensation and Gibbs-modeled vector fields,” in *Proc. IEEE IMDSP*, Mar. 1996, pp. 126–127.

[11] S. Sethuraman, M. W. Siegel, and A. G. Jordan, “A multiresolution region based segmentation scheme for stereoscopic image sequence compression,” in *Proc. SPIE EI-Digital Video Compression: Algorithms and Technologies*, vol. 2419, Feb. 1995, pp. 265–274.

[12] S. Malassiotis and M. G. Strinzis, “Object-based coding of stereo image sequences using 3-D model,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 892–905, Dec. 1997.

[13] H. Watanabe and S. Singhal, “Windowed motion compensation,” in *Proc. SPIE VCIP*, vol. 1605, Nov. 1991, pp. 582–589.

[14] M. Orchard and G. Sullivan, “Overlapped block motion compensation: An estimation-theoretic approach,” *IEEE Trans. Image Processing*, vol. 5, pp. 693–699, Mar. 1994.

[15] H. Dankworth and W. Niehsen, “Analytical overlapped block motion compensation,” in *Proc. NORSIG97*, 1997, pp. 75–80.

[16] R. Rajagopalan, E. Feig, and M. T. Orchard, “Motion optimization of ordered blocks for overlapped block motion compensation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 119–123, Apr. 1998.

[17] S. Nogaki and M. Ohta, “An overlapped block motion compensation for high quality motion picture coding,” in *Proc. IEEE ISCAS*, 1992.

[18] J. K. Su and R. M. Mersereau, “Noniterative rate-constrained motion estimation for OBMC,” in *Proc. IEEE ICIP*, 1997, pp. 33–36.

[19] T. Kuo and C. Kuo, “Motion optimization of ordered blocks for overlapped block motion compensation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 119–123, Apr. 1998.

[20] B. Tao and M. T. Orchard, “Joint application of overlapped block motion compensation and loop fieltering for low bit-rate video coding,” in *Proc. IEEE ICIP*, 1997.

[21] J. E. Besag, “Spatial interaction and the statistical analysis of lattice systems,” *J. Roy. Statist. Soc.*, vol. B36, pp. 192–236, 1974.

[22] W. Woo and A. Ortega, “Modified overlapped block matching for stereo image coding,” in *Proc. SPIE EI-VCIP*, vol. 3653, Jan. 1999.

[23] B. Girod, “Motion-compensating prediction with fractional-pel accuracy,” *IEEE Trans. Commun.*, vol. 41, pp. 604–611, Apr. 1993.

[24] K. Illgner and F. Muller, “Analytical analysis of subpel motion compensation,” in *Proc. PCS*, 1997, pp. 135–140.

[25] W. Woo and A. Ortega, “Dependent quantization for stereo image coding,” in *Proc. SPIE EI-VCIP*, vol. 3309, Jan. 1998, pp. 902–913.

[26] ——, “Blockwise dependent bit allocation for stereo image coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 861–867, Sept. 1999.