Journal of
Electronic Imaging

SPIEDigitalLibrary.org/jei

Reduced-complexity multiview prediction scheme with content-adaptive disparity vector estimation

Aykut Avci
Jan De Cock
Jelle De Smet
Youri Meuret
Peter Lambert
Herbert De Smet

SPIE
IS&T
imaging.org

Reduced-complexity multiview prediction scheme with content-adaptive disparity vector estimation

Aykut Avci
Ghent University
Department of Electronics and Information Systems
Centre for Microsystems Technology
Technologiepark 914
9052 Ghent, Belgium
E-mail: aykut.avci@ugent.be

Jan De Cock
Ghent University-IBBT
Department of Electronics and Information Systems
Multimedia Laboratory
Gaston Crommenlaan 8 bus 201
B-9050 Ledeberg-Ghent, Belgium

Jelle De Smet
Ghent University
Department of Electronics and Information Systems
Centre for Microsystems Technology
Technologiepark 914
9052 Ghent, Belgium

Youri Meuret
Vrije Universiteit Brussel
Department of Applied Physics and Photonics
Brussels Photonics Team Pleinlaan 2
1050 Brussels, Belgium

Peter Lambert
Ghent University-IBBT
Department of Electronics and Information Systems
Multimedia Laboratory
Gaston Crommenlaan 8 bus 201
B-9050 Ledeberg-Ghent, Belgium

Herbert De Smet
Ghent University
Department of Electronics and Information Systems
Centre for Microsystems Technology
Technologiepark 914
9052 Ghent, Belgium
and
imec-Cmst
Technologiepark 914
9052 Ghent, Belgium

Paper 12017P received Jan. 13, 2012; revised manuscript received May 15, 2012; accepted for publication Jun. 14, 2012; published online Aug. 22, 2012.

0091-3286/2012/$25.00

Abstract. *Disparity estimation is a highly complex and time consuming process in multiview video encoders. Since multiple views taken from a two-dimensional camera array need to be coded at every time instance, the complexity of the encoder plays an important role besides its rate-distortion performance. In previous papers we have introduced a new frame type called the D (derived) frame that exploits the strong geometrical correspondence between views, thereby reducing the complexity of the encoder. By employing the D frames instead of some of the P frames in the prediction structure, significant complexity gain can be achieved if the threshold value, which is a keystone element to adjust the complexity at the cost of quality and/or bit-rate, is selected wisely. A new adaptive method to calculate the threshold value automatically from existing information during the encoding process is presented. In this method, the threshold values are generated for each block of each D frame to increase the accuracy. The algorithm is applied to several image sets and 20.6% complexity gain is achieved using the automatically generated threshold values without compromising quality or bit-rate.* [DOI: 10.1117/1.JEI.21.3.033009]

1 Introduction

Recently, there has been a growing interest in three-dimensional (3-D) display technologies.[1–3] As these technologies become more common, improving the realism of the reproduced scene, which is highly dependent on then number of available views, becomes more relevant. The view images can be captured from different viewpoints of a scene by using a two-dimensional (2-D) camera array. A smoother transition between views can be obtained by increasing the number of cameras located in the 2-D camera array. However, this comes at the price of an increased amount of image data which needs to be encoded to store and transmit the data efficiently.

The captured multiview videos (MVV) for different views can be encoded separately by a state-of-the-art video codec like H.264/AVC (advanced video coding), which is called simulcast coding. Although coding each video individually is an easy option to solve the problem, it is not the most efficient approach since the inter-view correlations between views are overlooked. However, exploiting the inter-view correlations between views in a MVV is a very challenging task due to the high computational complexity which makes it difficult to implement. Recently, some methods have been proposed in order to reduce the complexity of the encoder. Kannangara et al. reported that the complexity reduction can be obtained by decreasing or early termination of the number of prediction modes for a macroblock.[4] Choi et al. proposed an early SKIP mode decision and selective intra-mode decision methods to reduce the computational complexity of the encoder.[5] Yin et al. employed an algorithm which jointly optimizes the motion estimation and the mode decision process.[6] Deng et al. introduced an iterative search method for motion and disparity estimation by utilizing the stereo-motion consistency constraints for the stereoscopic video coding.[7] Many methods to reduce the enormous computational load to encode multiview videos have been proposed in the literature. For example, Li et al. proposed a method to reduce the motion and disparity estimation computing by limiting the search region.[8] Zeng et al. described an algorithm to achieve the complexity reduction by terminating the mode decision process if the rate-distortion (R-D) cost of a macroblock is lower than an adaptively calculated threshold value.[9] Ding et al. proposed a content-aware prediction algorithm with inter-view mode decision to reduce the high computational complexity of the multiview video encoder.[10] Shen et al. proposed a fast disparity and motion estimation algorithm based on the homogeneity which is estimated by looking at the motion vectors of neighboring and previously encoded corresponding macroblocks.[11,12] Zhu et al. described a fast disparity estimation algorithm which utilizes the spatio-temporal correlation and the temporal variation of the disparity field.[13] However, all methods mentioned above are achieving the complexity reduction in a MVV which contains either only horizontally distributed views or some selected views of view images captured by a 2-D camera array. To encode multiview image data captured by a 2-D camera array at one time instant, different prediction schemes have already been proposed. Merkle et al. proposed prediction schemes shown in Fig. 1 (MVCS-1 and MVCS-2) to encode a MVV sequence.[14] Although the redundancies in the view domain can be successfully exploited by one of those prediction schemes, the computational complexity of the encoder is high due to the complex and time consuming disparity vector searches. To reduce the existing complexity of the encoder, we proposed, in our previous work, a new frame type called D (derived) frame in which the disparity vector of a block can be derived from the other views due to the strong geometrical correspondence existing between adjacent views.[15–17] As a result of employing D frames instead of some of the P frames in the prediction scheme as shown in Fig. 2 (CR-MVCS), remarkable complexity gains can be achieved.

A threshold value in the algorithm determines for which blocks the geometrical correspondence is strong enough to justify derivation of their disparity vectors from other views. By wisely choosing the threshold value, the complexity reduction can be achieved without compromising the quality or bit-rate. In our previous work, the threshold values were determined experimentally at different QPs.[18] The arrows in Fig. 1 and Fig. 2 show the prediction direction of the frames and the frames labeled I and P are the standard-conforming intra-coded and predicted frames that are well known from H.264/AVC video coding standard.

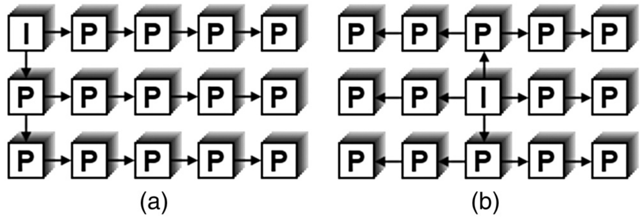


Fig. 1 Prediction schemes proposed by Merkle et al.[14] for 5×3 views. (a) =MVCS – 1, and (b) =MVCS – 2.

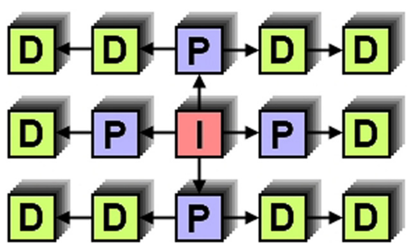


Fig. 2 Modified version of MVCS-2 which is a traditional multiview encoding scheme. Some of the P frames are replaced with D frames (CR-MVCS).

Since the threshold values are calculated experimentally in our previous work, they need to be signaled in the beginning of the encoding of any D frame in the prediction scheme. To overcome this extra information, an automated way of calculating the threshold values from previously encoded P frames at different QPs is presented in this work.

The remainder of this paper is organized as follows. The framework of the D frame and the proposed algorithm to automate the threshold values are described in Sec. 2. The proposed prediction scheme (CR-MVCS) using different weighing functions is compared with the traditional prediction scheme (MVCS-2) on various image sets in Sec. 3 and finally Sec. 4 concludes the paper.

2 Background

2.1 *D Frame*

The D frame is a version of the standard-conforming P frame where mainly the disparity estimation part is modified. The flow chart of the disparity estimation process to be run for each block of the D frame is shown in Fig. 3. The main idea of the D frame is that the complexity of the encoder can be reduced by skipping the disparity estimation process for some of the blocks in the D frame depending on the fidelity of their derived disparity vectors. The disparity vector of a block in a D frame can for most blocks be derived from one or more of the other views since there is strong geometrical correspondence between views in the 2-D camera array. The derived disparity vector needs to be checked for the convenience because the derivation process can fail in some blocks due to occlusions, illumination changes, or insufficient texture information. For this purpose, the R-D cost of the derived disparity vector should be checked against the threshold value which is obtained from the same block from which the disparity vector is derived. This will be discussed in more detail in the next section. If the derived disparity vector passes the convenience check successfully, the encoder skips the time consuming disparity estimation process for that block and the derived disparity vector is saved, which results in complexity gain for the encoder. In the other case, the disparity estimation is run as it is done in the P frame. At this point, it is clear that the threshold value plays a crucial role in the determination of convenience of the derived disparity vector. Since the D frame is a modified standard-conforming P frame, the encoder shows the same performance as the P frame if the threshold value is set equal to zero, meaning that none of the derived disparity vectors are used and the disparity estimation will be performed for all the blocks in the frame. As the threshold value increases beyond a certain value, the accuracy of the derived disparity vector relaxes. Also the quality and bit-rate of the encoded data start to degrade. It is this value of the threshold with which maximum complexity gain can be achieved. In this paper, a novel algorithm for the determination of this critical value of the threshold is proposed.

The arrows in the prediction scheme in Fig. 2 show the prediction direction of the frames and give no information about the derivation of the disparity vectors for the D frames. The disparity vectors of the D frames situated in the middle row are extrapolated from the disparity vectors of the adjacent P frame in the same row. The disparity vectors of the D frames in the other rows are derived from the disparity vectors of the corresponding middle row frame in the same column.

The prime purpose of the D frame is to reduce the encoder complexity that is due to the excessive amount of image data. Since no change has been made in the syntax of the encoder, the encoded bitstream is fully compliant to the H.264/AVC specification. In other words, no special decoder is needed to decode the encoded bitstream.

2.2 *Threshold Value Calculation and Automation*

As it is stated in the previous section, the R-D performance of the encoded data starts deteriorating if the threshold value rises above a certain value that depends on the image set and on QP. Therefore, the threshold values should be obtained during encoding by considering the nature of the image set and the QP. In order to increase the accuracy and the efficiency of the encoding process, the threshold value should be calculated block by block instead of applying a single threshold value to the whole prediction scheme, as we did in our previous work.[18,19]

Since the D frame is using the disparity vector information of the blocks in a P frame, the P frame must be encoded before we can begin to encode the D frame. Therefore, the previously encoded P frame can be employed to estimate the threshold values of the blocks during the encoding of the D frame. This can be done by extracting the R-D cost values of blocks in an already encoded P frame, which are to be used as the threshold values of the corresponding blocks in the D frame after a certain weighting process.

The weighting process is necessary to derive the threshold values from the extracted R-D cost values. A small R-D cost value of a block in a P frame, which for instance typically occurs in backgrounds, indicates a good match between the

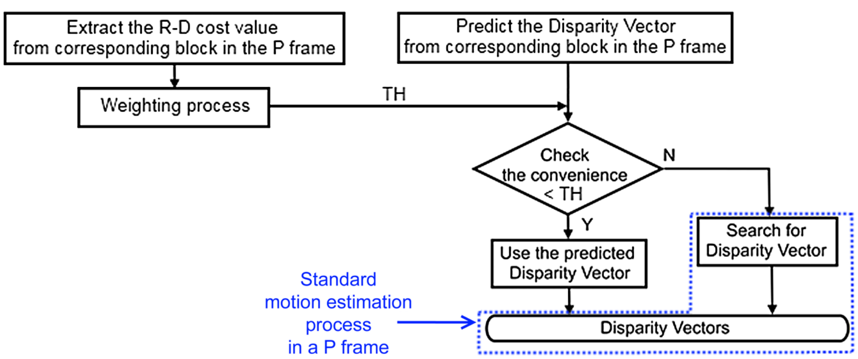


Fig. 3 The flow chart of the disparity estimation process of the D frame.

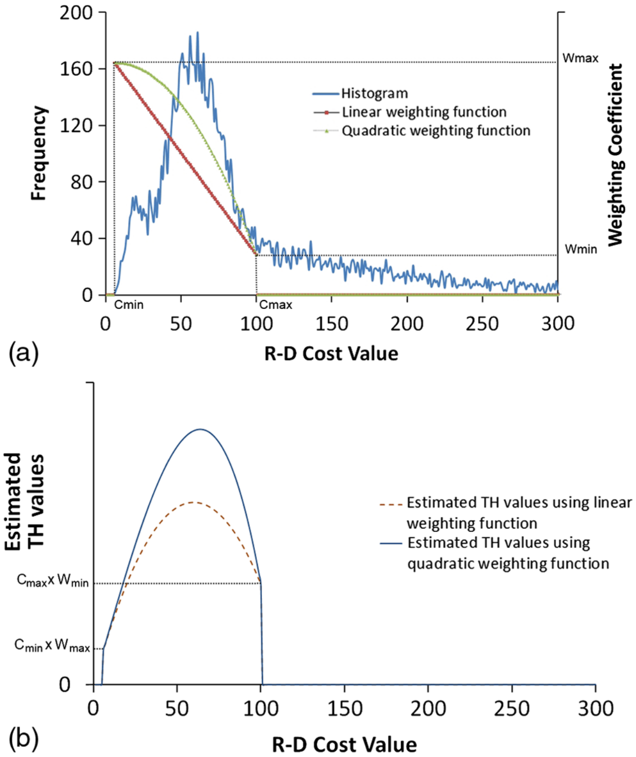


Fig. 4 (a) Histogram graph of the R-D cost values extracted from a P frame in the prediction scheme (left ordinate) and two different weighting functions (right ordinate). (b) Estimated threshold values corresponding to the histogram shown in (a).

reference block and the candidate block, confirming the success of the encoding. Therefore, this small cost value should be weighted by a large coefficient because the encoder is expected to find also a good match between the corresponding block in the D frame and its reference block. On the other hand, blocks with large R-D cost values, which for instance occur in occlusions, need to be weighted by a small coefficient to have smaller threshold values, in order to restrict possible deterioration caused by derived disparity vectors that point to irrelevant reference blocks.

Two different weighting functions which underline the fact that there is an inverse relation between the R-D cost values and the weighting coefficients are shown in Fig. 4(a). The histogram graph of the extracted R-D cost values from the P frame is also plotted to explain the weighting process in a better way. While the left ordinate presents the frequency of the histogram graph, the right ordinate indicates the weighting coefficient. The starting and finishing points where the weighting functions are applied are marked as $C_{\min}$ and $C_{\max}$ in the x-axis which represents the R-D cost values. The estimated threshold values corresponding to the histogram graph shown in Fig. 4(a) is plotted in Fig. 4(b).

The linear weighting function (LWF) can be described as;

$$f(x) = \begin{cases} 0, & x < C_{\min} \text{ or } x > C_{\max} \\ \frac{W_{\max}(C_{\max}-x)+W_{\min}(x-C_{\min})}{(C_{\max}-C_{\min})}, & C_{\min} \le x \le C_{\max} \end{cases}. \quad (1)$$

The quadratic weighting function (QWF) can be described as;

$$f(x) = \begin{cases} 0, & x < C_{\min} \text{ or } x > C_{\max} \\ W_{\max} + \frac{W_{\min}-W_{\max}}{(C_{\max}-C_{\min})^2}(x-C_{\min})^2, & C_{\min} \le x \le C_{\max} \end{cases}, \quad (2)$$

where x is the extracted R-D cost value of a block from which the disparity vector is derived.

The threshold value can then be calculated by,

$$TH = x \times f(x). \quad (3)$$

The encoder runs the disparity estimation process for those blocks where the weighting function yields zero, which means that the R-D cost value of those blocks does not fall in between $C_{\min}$ and $C_{\max}$. For the blocks whose R-D cost values lie between $C_{\min}$ and $C_{\max}$, both weighting functions are applied to produce higher weighting coefficients for small R-D cost values and lower weighting coefficients for high R-D cost values.

3 Experimental Results

The prediction scheme with D frames (CR-MVCS) and the traditional prediction scheme (MVCS-2) are applied to different image sets released by the Stanford University Computer Graphics Laboratory.[20] These image sets contain various real world objects and they are in the form of rectified and cropped high resolution image sequences generated by a 2-D 17×17 camera grid: "Amethyst" (768×1024), "The Stanford Bunny" (1024×1024), "Eucalyptus Flowers," (1280×1536), "Treasure Chest" (1536×1280), and "Truck" (1280×960). The numbers between brackets represent the resolution in pixels of the respective image sets. There are some low resolution image sets provided by other sources in the literature. Since the aim of the method is to achieve complexity reduction while reaching the best R-D performance especially for high resolution images, those image sets are not considered for the current study. For the encoder, the standard-conforming JSVM 9.19.7.[21] is used and only 8×8 macroblock partitioning is allowed. The reason for using JSVM reference software instead of JMVM is that the JMVM reference software does not enable us to encode the prediction schemes proposed in this paper. Since both software packages are derived from the same code base for H.264/AVC, and the JSVM software is more flexible to implement the idea in this paper, the JSVM software is chosen as encoder in this work.

In order to study the influence of experimentally determined values of $C_{\max}$ and $W_{\max}$ and to calculate the difference between CR-MVCS and MVCS-2, we employ the Bjontegaard method which produces two equivalent results for quality and bit-rate difference (BD-PSNR and BD-Rate, respectively).[22] Since it is proven in our previous work that MVCS-2 is exhibiting better R-D performance than MVCS-1 for the current image sets, MVCS-1 is not included in the experiments in this work.

3.1 *Value Determination*

The $C_{\max}$ and $C_{\min}$ parameter values used to construct the weighting functions are dynamically calculated during the encoding process of the D frame by using the extracted R-D cost values of the blocks from the P frame from which the disparity vectors are derived. While the $C_{\min}$

Table 1 W_{max} parameter values for different QPs.

	W_{max}	
QP	**Linear weighting function**	**Quadratic weighting function**
22	0.9	0.7
27	1.6	1.25
32	1.8	1.5
37	2.3	1.7

value corresponds with the minimum R-D cost value, C_{max} is chosen equal to the average of all the R-D cost values from the reference P frame, which is found experimentally. The parameter corresponding with the minimum weighting coefficient, denoted W_{min}, is set equal to zero in this experiment in order to have a continuous function. The optimum value of W_{max} is determined for different QPs as listed in Table 1, which are calculated by observing the BD difference and the complexity results while changing the W_{max} parameter value.

In order to verify the accuracy of the experimentally determined values of C_{max} and W_{max} as mentioned above, various data points lying within ±60% of the respective experimental values are tested on the Amethyst image set. Figure 5(a) and 5(b) shows the BD-PSNR and BD-Bit-Rate performances between MVCS-2 and CR-MVCS where C_{max} and W_{max} parameters are used to construct the weighting functions in terms of quality and bit-rate, respectively. The complexity results which are obtained while generating the results in Fig. 5(a) and 5(b) are plotted in Fig. 5(c). In order to draw a conclusion on the effectiveness of C_{max} and W_{max}, all three figures in Fig. 5 need to be interpreted together. On the one hand, it can be concluded from Fig. 5(a) and 5(b) that the encoded bitstream shows no R-D difference (up to four decimal places) compared to MVCS-2 if the C_{max} and W_{max} are set equal or lower than their experimentally determined values which correspond to zero in the x-axis in Fig 5. On the other hand, however, the complexity of the encoder increases as the C_{max} and W_{max} decreases [Fig. 5(c)]. Therefore, experimentally determined values of C_{max} and W_{max} yield maximum complexity gain while preserving R-D performance of the encoder.

The experiment shows parallel results when the method is applied to any of the other image sets. Note that the outcomes of the plotted graphs in Fig. 5 where the x-axis is equal to zero can also be observed in Table 2 where the comparative results of all image sets with experimentally determined values of C_{max}, C_{min}, W_{max}, and W_{min} are presented.

3.2 *Comparative Results*

Multiview images from one time instant of the image sets listed above are encoded with the MVCS-2 and the CR-MVCS schemes, utilizing either the linear or the quadratic weighting function for several QP values such as 22, 27, 32, and 37. In addition to the bitrate and quality results of encoding, the number of blocks for which the disparity estimation process is run is also recorded in order to compare the

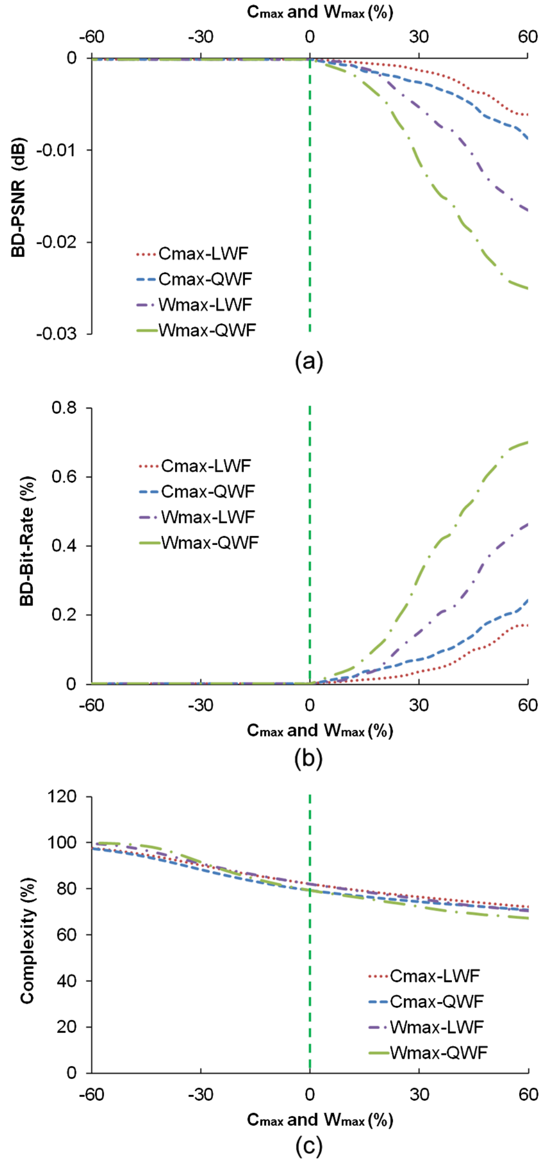


Fig. 5 Impact of C_{max} and W_{max} on the Bjontegaard quality, bit-rate and, complexity of the encoder. In the abscissa of these graphs, the coefficients are varied over ±60% with respect to the experimentally determined values.

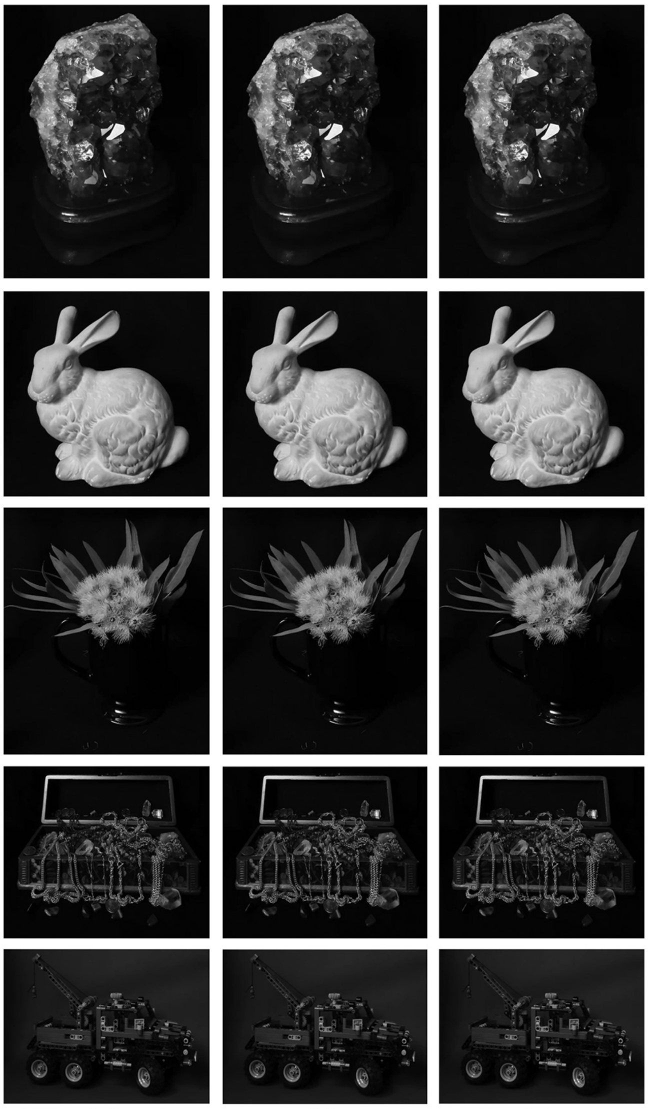
prediction schemes and the weighting functions in terms of complexity. Since the disparity estimation process has to be performed for all blocks in the P frames of the MVCS-2 prediction scheme, the complexity of the MVCS-2 is 100% regardless of the QP. However, the complexity varies for the CR-MVCS prediction scheme because of skipped disparity estimation processes. In order to show the efficiency of the CR-MVCS with weighting functions, the number of searched blocks of all the frames are calculated for different QPs and averaged. As it can be concluded from the results shown in Table 2, the CR-MVCS prediction scheme utilizing either linear or quadratic weighting function is showing the same R-D performance as the MVCS-2 prediction scheme since the BD-Bit-Rate and BD-PSNR values are equal to zero. However, the CR-MVCS prediction scheme requires lower complexity to achieve the same R-D results, which is the main advantage of having the D frames instead of P frames in the prediction scheme. As can been seen

Table 2 Experimental results.

	Avg. searched blocks [%]		BD-rate [%]		BD-PSNR [dB]	
Image set	**Linear**	**Quadratic**	**Linear**	**Quadratic**	**Linear**	**Quadratic**
Amethyst	82.9	79.4	0.0	0.0	0.0	0.0
The S. Bunny	90.5	86.8	0.0	0.0	0.0	0.0
Euc. Flowers	88.4	84.5	0.0	0.0	0.0	0.0
Tre. Chest	81.3	80.0	0.0	0.0	0.0	0.0
Truck	92.2	88.4	0.0	0.0	0.0	0.0

from Table 2 that in the case of the Amethyst image set, only 79.4% of all the blocks were searched for disparity vector, which implies the complexity gain of 20.6% for the overall disparity estimation process of the encoder. It is also clear from Table 2 that the quadratic weighting function yields better complexity results compared to the linear weighting function. The quadratic weighting function is more efficient because the value of the produced coefficient is greater than or equal to the one obtained from a linear weighting function. This results in a higher threshold value which enables us to skip relatively larger number of blocks yielding higher complexity efficiency. On the other hand, there exists a trade-off between the complexity gain and R-D performance. For a very high value of the threshold the encoder will run the disparity estimation process for fewer blocks but the R-D performance will degrade. However, it should be noted that the weighting functions are not limited to the two implementations which are presented in this work and different weighting functions can be designed. In this case, appropriate parameters to construct the weighting function should also be investigated.

View images encoded with the CR-MVCS and MVCS-2 prediction schemes were also compared visually and did not show any differences in quality. In Fig. 6, the images coming from the top right camera view of all image sets are shown after encoding and subsequent extraction. The MVCS-2 prediction scheme, which comprises only I and P frames, is compared with the CR-MVCS prediction scheme for both the linear and the quadratic weighting function, constructed using the experimentally determined values of $C_{\min}$ and $W_{\max}$.



(a) (b) (c)

Fig. 6 Decoded image samples of top-right view in the image set with experimentally determined values of $C_{\min}$, $W_{\max}$, and $\mathrm{QP} = 32$. (a) Encoded with MVCS-2. (b) Encoded with CR-MVCS by employing a linear weighting function. (c) Encoded with CR-MVCS by employing a quadratic weighting function.

4 Conclusion

Complexity is an important issue for multiview video encoders since multiple views need to be coded for every time instant. A modified type of P frame, named D frame, which takes advantage of the strong geometrical relationship among views was introduced in our previous work. By employing D frames instead of P frames in the prediction schemes, the complexity of the encoder can be decreased drastically if the threshold value is selected wisely.

In this paper, an algorithm to determine the threshold values automatically for different QPs is presented. The algorithm enables the calculation of the threshold values for each block independently in a D frame, which makes it more accurate and efficient. The main idea behind the algorithm is that the R-D cost values of blocks of previously encoded

P frames can be employed to determine the threshold values of the blocks in the D frame. In other words, the threshold value of a block in a D frame can be calculated from the R-D cost of the corresponding block in the P frame from which the disparity vector is derived. Since there is an inverse relation between the R-D cost and the threshold value, a weighting process needs to be applied on the R-D cost value before it can be used as a threshold value in a D frame.

Two different weighting functions named linear and quadratic weighting functions are presented in this work. The weighting functions are constructed in such a way that they produce a higher weighing coefficient for small R-D cost values and lower weighting coefficients for high R-D cost values. During the encoding of the D frame, the R-D cost of the corresponding block in the P frame is weighted by the coefficient calculated by means of the weighting function.

According to the experimental results, the proposed prediction scheme with the D frames shows the same R-D performance as the traditional prediction scheme with P frames, and this holds true for both proposed weighting functions. However, the proposed prediction scheme allows to skip the disparity estimation process for some of the blocks, which results in a net complexity gain that is most pronounced in the case of the quadratic weighting function.

The results can further be improved by implementing other algorithms to achieve complexity reduction for the blocks for which the disparity estimation process is run (blocks whose derived disparity vector is nonconvenient) as it is done in a regular P frame.

Acknowledgments

This work is supported by the Research Foundation-Flanders (FWO-Vlaanderen). The project is titled “Compact LCOS projection displays for high-quality 3-D images with high spatial and angular resolution.”

References

1. N. A. Dodgson, “Autostereoscopic 3D displays,” *Computer* **38**(8), 31–36 (2005).
2. P. Benzie et al., “A survey of 3DTV displays: techniques and technologies,” *IEEE Trans. Circuits Syst. Video Technol.* **17**(11), 1647–1658 (2007).
3. N. Holliman et al., “Three-dimensional displays: a review and applications analysis,” *IEEE Trans. Broadcast.* **57**(2), 362–371 (2011).
4. C. S. Kannangara et al., “Low complexity skip prediction for H.264 through Lagrangian cost estimation,” *IEEE Trans. Circuits Syst. Video Technol.* **16**(2), 202208 (2006).
5. I. Choi, J. Lee, and B. Jeon, “Fast coding mode selection with rate-distortion optimization for MPEG-4 part-10 AVC/H.264,” *IEEE Trans. Circuits Syst. Video Technol.* **16**(12), 1557–1561 (2006).
6. P. Yin et al., “Fast mode decision and motion estimation for JVT/H.264,” in *Proc. of International Conference on Image Processing (ICIP)*, pp. 853–856, IEEE, Piscataway, NJ (2003).
7. Y. L. Chan et al., “Fast motion and disparity estimation with adaptive search range adjustment in stereoscopic video coding,” *IEEE Trans. Broadcast.* **58**(1), 24–33 (2012).
8. X. Li et al., “Fast disparity and motion estimation based on correlations for multiview video coding,” *IEEE Trans. Consum. Electron.* **54**(4), 2037–2044 (2008).
9. H. Zeng, K. K. Ma, and C. Cai, “Mode-correlation-based early termination mode decision for multi-view video coding,” in *Image Processing (ICIP)*, pp. 3405–3408, IEEE, Piscataway, NJ (2010).
10. L. F. Ding et al., “Content-aware prediction algorithm with inter-view mode decision for multiview video coding,” *IEEE Trans. Multimedia* **10**(8), 1553–1564 (2008).
11. L. Shen et al., “Selective disparity estimation and variable size motion estimation based on motion homogeneity for multi-view coding,” *IEEE Trans. Broadcast.* **55**(4), 761–766 (2009).
12. L. Shen et al., “View-adaptive motion estimation and disparity estimation for low complexity multiview video coding,” *IEEE Trans. Circuits Syst. Video Technol.* **20**(6), 925–930 (2010).
13. W. Zhu et al., “Fast disparity estimation using spatio-temporal correlation of disparity field for multiview video coding,” *IEEE Trans. on Cons. Elec.* **56**(2), 957–964 (2010).
14. P. Merkle et al., “Efficient prediction structures for multiview video coding,” *IEEE Trans. Circuits Syst. Video Technol.* **17**(11), 1461–1473(2007).
15. A. Avci et al., “Efficient disparity vector coding for multi-view 3D displays,” *Three-Dimensional Image Process. (3DIP) Appl.* **7526**(1), 752609 (2010).
16. Y. Kim et al., “Efficient disparity vector coding for multiview sequences,” *Signal Process. Image Commun.* **19**(6), 539–553 (2004).
17. T. Y. Chung et al., “Multi-view video coding with view interpolation prediction for 2D camera arrays,” *J. Vis. Commun. Image Represent.* **21**(5–6), 474–486 (2010).
18. A. Avci et al., “Efficient disparity vector prediction schemes with modified P frame for 2D camera arrays,” *J. Vis. Commun. Image Represent.* **23**(2), 287–292 (2012).
19. A. Avci et al., “Reduced complexity multi-view video coding scheme for 2d camera arrays,” in *3DTV Conference: The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1–4, IEEE, Piscataway, NJ (2011).
20. Stanford University, “The Stanford light field archive,” http://lightfield.stanford.edu.
21. ITU-T Recommendation H.264, ISO/IEC 14496-10 Advanced Video Coding, “Advanced Video Coding Generic Audiovisual Services,” (March 2009).
22. G. Bjontegaard, “Calculation of average PSNR differences between RD-curves,” *ITU-T SG16/Q.6 Doc. VCEG-M33*, Austin, TX (2001).

Biographies and photographs of the authors are not available.