Spatially Scalable Video Coding with In-band Prediction

Xin Jin[*a], Xiaoyan Sun[b], Feng Wu[b], Guangxi Zhu[a]

[a]Dept. of Electronics & Information, HUST, Wuhan, P.R.China 430074
[b]Microsoft Research Asia, Beijing, P.R.China 100080

ABSTRACT

The discrete wavelet transform (DWT) has been widely used in scalable video coding for its advantages in multi-resolution analysis and subband decomposition. In this paper, a spatially scalable video coding system based on H.264 coding method and in-band overcomplete discrete wavelet transform (ODWT) technique is proposed, which integrates the good compression performance of H.264 in low frequency domain with the efficient motion estimation of in-band ODWT in wavelet domain. Intra prediction, coefficients scan manner and inter prediction are improved to overcome the inefficiency of H.264 coding in high frequency subbands caused by different pixels distribution properties. Through series of subband analysis and statistical data retrieval for the three high frequency decompositions, intra prediction directions are optimized and three subsets of prediction mode are presented for the three high subbands respectively. They save over 30% bits for intra mode with similar performance. Moreover, novel zigzag scan tables are proposed to improve the coding efficiency by utilizing the oriented frequency characteristics of each high band. To inter prediction, an adaptive motion estimation method is proposed in which the motion information of low band is adaptively and effectively utilized to achieve much more accurate motion vector and more efficient motion compensation in high bands coding. Experimental results show that, all of the proposed methods endue the spatially scalable video coding system with over 0.4 dB gain in PSNR and 10.4% in rate reduction.

Keywords: Spatial scalability, subband analysis, adaptive motion estimation, video coding

1. INTRODUCTION

With the development of multimedia applications in heterogeneous networks, the scalable video coding systems draw great attention in both industry and research areas. Within temporally, spatially and SNR scalable manner, spatially scalable video coding (SSVC) is of particular interest because of its broad application for a single video bitstream adapting to variable receiver display resolution capabilities and unstable transmission bandwidth. To the spatially scalable video stream, the base layer corresponds to the video bitstream at reduced spatial resolution and can be decoded individually, while the enhancement layers represent the textured information for reconstructing signals at higher resolutions.

Because of its outstanding performance in hierarchically signal decomposition, wavelet transforms have been widely used in multi-resolution analysis and many attempts of SSVC system[1-6]. These schemes could be divided into two categories: one is the 3-D wavelet video-coders with motion estimation (ME) and motion compensation (MC) in spatial domain[1,2]; the other is the 3D or in-band wavelet video-coders with ME and MC in the wavelet domain[3-6]. While in the second category, the in-band wavelet video approaches based on the classical hybrid coding scheme (such as MPEG2 and MPEG4) are widely used to benefit from the advantages of traditional coding methods with much better performance in low resolutions encoding when comparing with the first category approaches. Moreover, researches in [5] have proved that, for spatially scalable coding, the in-band motion compensation based on overcomplete wavelet transform (ODWT) can yield similar performance to the full resolution MC approaches under comparable MC accuracy. So, the in-band wavelet video system based on hybrid coding and ODWT is a desirable alternative for SSVC.

In order to integrate the good compression performance of H.264[7] in low frequency domain with the attractive advantages of in-band ODWT in wavelet domain ME and MC, an approach of SSVC system is proposed in this paper based on the two techniques to provide the desirable spatial scalability while maintaining the good coding efficiency. It is proved that the H.264 coding scheme can work well in low frequency subband encoding. However, it is hard to

---

[*]Contact goldcamel2001@yahoo.com; phone 86 27 87543482; http://www.bbcw.org; Dept. of Electronics & Information, Huazhong University of Science & Technology, Wuhan, P.R.China 430074
This work has been done while the author is with Microsoft Research Asia

perform so efficiently, especially on the prediction and the coefficients coding, in high frequency domain because of the different pixels distribution properties. Though the ODWT with low-band-shift (LBS)[8] method is applied in the in-band ME and MC for our system, the efficiency of prediction and motion estimation in high frequency subbands still shows kind of insufficient because of the differences between the high frequency domain and the low frequency. The goal of this paper is to propose several technical improvements in intra prediction, reordering scan and motion estimation for high frequency decompositions. All the researches are based on detailed analysis and statistical data retrieval on directional and motional correlation among those decomposed subbands. Experimental results show that all of these techniques endure our proposed SSVC system with higher performance and better adaptability in balance of complexity.

In this paper, the architecture of the proposed spatially scalable video coding system is presented in section 2 and the technical improvements are described in detail in section 3, focusing on intra prediction, reordering scan and motion estimation with statistical data retrieving, characteristic analyzing, performance optimizing and complexity balancing. The performance of the proposed scheme is evaluated in section 4. Finally, section 5 concludes the paper.

2. SYSTEM ARCHITECTURE

Figure 1 shows the architecture of the proposed in-band spatially scalable video coding system based on H.264 hybrid coding and overcomplete motion estimation and motion compensation.

As shown in the figure, every image of the input video is firstly decomposed into one low frequency subband, low-low (LL) band, for low resolution base layer and three high frequency subbands, low-high (LH), high-low (HL) and high-high (HH) band, with detailed and textured information for enhancement layer. Then the four-subband “images” of wavelet coefficients are sent to H.264 codec sequentially for block-based predicting and encoding. Both intra and inter prediction are enabled for each subband, which are determined by rate distortion optimization. The block-based ME and MC are performed in wavelet domain. Different from common solutions, half-pixel interpolation in the spatial domain is firstly performed on the original full resolution image which is reconstructed by inverse DWT. Then, a reference subband conversion from complete to overcomplete is applied using ODWT with LBS proposed in [8] to overcome the shift-variant property during the conversion. Higher efficiency of ME and MC is achieved by in-band block-matching in these four overcomplete reference subbands. After inter and intra predictions, the residues of the four subbands are coded using H.264 4x4 block-based transform and quantization. Finally, the context-adaptive binary arithmetic coding (CABAC) is used to improve the efficiency of entropy coding for quantized residues to generate the bitstreams. The adaptive context model of CABAC permits adaptation to nonstationary symbol statistics to match the variant distribution of high frequency subband coefficients[9].

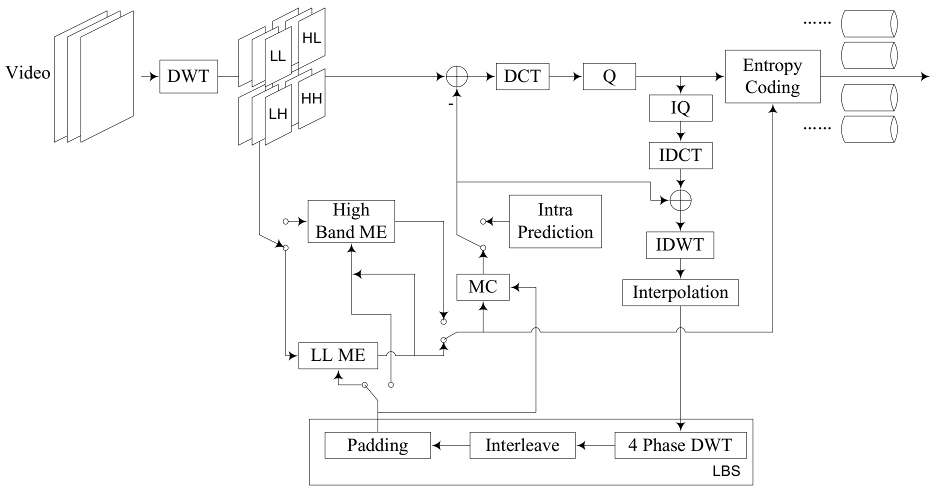


Figure 1: System architecture of the proposed spatially scalable video coding.

In the proposed system, good compression performance of H.264 is integrated with attractive advantages of in-band ODWT for coding efficiency of spatial scalability. Since the LL band shows similar features to nature video sequences, the coding scheme of H.264 is able to work well in low frequency subband. In other words, the base layer of the proposed SSVC scheme can be coded efficiently. However, the motion fields formed by in-band ME and MC of high frequency subbands are not so smooth as the ones of the LL band, thus the block-based prediction method of H.264 can not work so efficiently as that in low frequency domain. By subband comparing and analyzing, the difference of pixels distribution properties between the high frequency and the low frequency decompositions is observed to result in this compression inefficiency. Consequently, novel techniques are proposed for intra prediction, reordering scan manner and motion estimation to improve the system performance in high frequency domain. The balance between performance optimization and complexity is also discussed for the system realization.

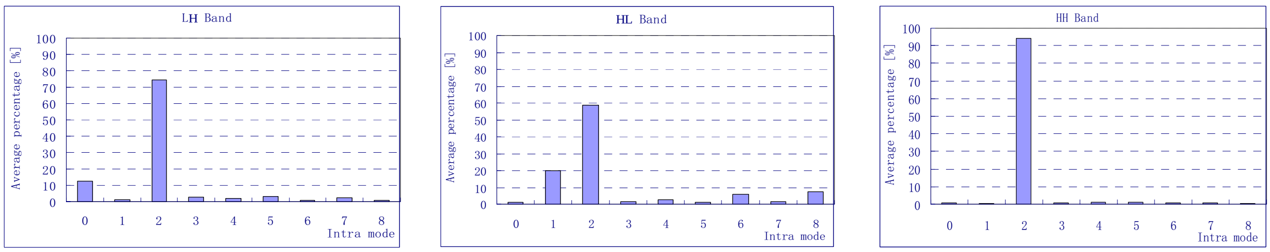
3. TECHNOLOGY IMPROVEMENT

In this section, several techniques optimizations are described focusing on intra prediction, reordering scan manners and ME based on the subband analyses. The test results are presented in each subsection to demonstrate the performance improvements.

3.1 Optimization of intra prediction

Investigating the directional relativity inside and among the neighboring blocks, H.264 uses 9 directions in the intra prediction[7], which are suitable to predict directional structures in a picture such as edges at various angles[9]. It evidently improves the coding efficiency of intra-coded blocks. While in the high frequency domain, the dependency inside and among blocks needs further discussion. The wavelet decomposition can be interpreted as signal decomposition in a set of independent, spatially oriented frequency channels. After wavelet transform, LH band represents horizontal high frequencies, HL band represents vertical high frequencies and HH band represents high frequencies in both directions[10]. Respectively, they show the vertical, the horizontal and the corner edges of one frame. Consequently, it is reasonable to optimize the intra prediction in high frequency subbands by fully utilizing the directional information and still maintain its efficiency with lower complexity and lower payload.

In order to study the directional properties of the three high frequency decompositions and verify the feasibility of the simplification in intra prediction, the selecting frequency of 9 intra prediction modes in high frequency subbands are figured out firstly based on rate distortion optimization and the average usage of each mode is calculated secondly. In order to make the results more reliable, various sequences were tested covering the vast moving properties as Football, the detailed anomalous moving properties as Bus, the obvious textured properties as Foreman, the camera moving and focus varying properties as Coastguard and the smooth moving properties as Container. Every four QP from 16 to 36 were tested in full length of every sequence using two kinds of group of picture (GOP) style as entire intra coded sequence (I-I-I-…) and the first frame intra coded sequence (I-P-P-P-…). The average statistical usage of the nine intra prediction modes (numbering as [7] defined) in the three high frequency subbands is shown in Figure 2.



(a) (b) (c)

Figure 2: Distribution of nine intra prediction modes in: (a) LH band; (b) HL band and (c) HH band.

As shown in the figure, besides the DC prediction, the usage of some intra prediction directions are dominant in the intra prediction, which varying with the subband and matching the spatially oriented properties of high frequency decompositions described above very closely. So, the one set of 9 intra prediction directions can be reduced to three subsets of 5 directions for the three high bands correspondingly for the intra prediction. As shown in Figure 3, four directions with DC prediction are applied to the three high bands respectively, which altogether occupy above 95% of usage in the previous counting with 30% bits saving in high band intra mode coding.

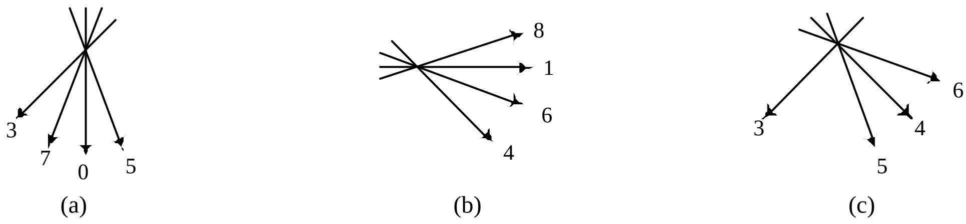


Figure 3: Optimized intra prediction directions for: (a) LH band; (b) HL band and (c) HH band.

By the optimization of intra prediction, the complexity of intra prediction is reduced to 2/3 by comparing with the original method, while maintaining the similar performance. To meet the requirement on even lower complexity, the intra prediction scheme could be further simplified by reducing the intra prediction mode in HH band to only one mode of DC prediction. Because of the stochastic distribution characteristic of HH band coefficients, DC prediction works much more efficient in HH band than in others. In this case, no payload on intra prediction mode is needed in HH band and the complexity of intra prediction is reduced to nearly 1/2 with slightly loss in performance.

3.2 Optimization of reordering scan

Quantized transform coefficients are required to be encoded as compactly as possible for transmission and storage. So, in H.264, coefficients reordering scan, named zigzag scan, is applied after quantization to group together nonzero coefficients and enable efficient representation of the remaining zero-valued quantized coefficients. The optimum scan order is determined by the distribution of nonzero DCT coefficients. In the low frequency domain, nonzero coefficients after DCT and quantization are concentrated into a small number of values around the top-left (0,0) position roughly symmetrically because of the energy compaction effect of DCT. Thus the suitable scan order is a zigzag starting from the top-left position[11]. While in the high frequency domain, the distribution is kind of different.

The directional correlation is reserved among residuals after block-based MC prediction or intra prediction in high frequency subbands. Because of the characteristic of orthogonal transform, the oriented properties of high bands are still preserved in coefficients after DCT and quantization. So, in order to group the nonzero coefficients together at the start of the reordered array, the distribution property of nonzero quantized coefficients in the three high bands was investigated. By observing the frequency of nonzero coefficients after DCT and quantization for every position of 4×4 block, three optimal scan paths which are suitable for the distribution property of AC coefficients of Intra_16×16 mode (INTRA AC), DC coefficients of Intra_16×16 mode (INTRA DC) and quantized results of other prediction modes (4×4 AC) in the high bands are proposed respectively, as shown in Table1 to Table 3.

Table 1. INTRA AC, INTRA DC and 4×4 AC zigzag scan tables of LH band.

Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
INTRA AC	$C_{0,0}$	$C_{3,0}$	$C_{3,1}$	$C_{3,3}$	$C_{3,2}$	$C_{2,0}$	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{1,0}$	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{0,1}$	$C_{0,2}$	$C_{0,3}$
INTRA DC	$C_{0,0}$	$C_{3,0}$	$C_{2,0}$	$C_{3,1}$	$C_{2,1}$	$C_{3,2}$	$C_{3,3}$	$C_{1,0}$	$C_{2,2}$	$C_{2,3}$	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{0,1}$	$C_{0,2}$	$C_{0,3}$
4×4 AC	$C_{3,0}$	$C_{2,0}$	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{0,0}$	$C_{1,0}$	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{1,1}$	$C_{1,2}$	$C_{0,1}$	$C_{1,3}$	$C_{0,2}$	$C_{0,3}$

Table 2. INTRA AC, INTRA DC and 4×4 AC zigzag scan tables of HL band.

Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
INTRA AC	$C_{0,0}$	$C_{0,3}$	$C_{0,2}$	$C_{1,3}$	$C_{0,1}$	$C_{1,2}$	$C_{2,3}$	$C_{3,3}$	$C_{2,2}$	$C_{1,1}$	$C_{3,2}$	$C_{2,1}$	$C_{3,1}$	$C_{1,0}$	$C_{2,0}$	$C_{3,0}$
INTRA DC	$C_{0,0}$	$C_{0,3}$	$C_{0,2}$	$C_{1,3}$	$C_{0,1}$	$C_{1,2}$	$C_{2,3}$	$C_{1,1}$	$C_{2,2}$	$C_{3,3}$	$C_{3,2}$	$C_{1,0}$	$C_{2,1}$	$C_{3,1}$	$C_{2,0}$	$C_{3,0}$
4×4 AC	$C_{0,3}$	$C_{0,0}$	$C_{0,2}$	$C_{0,1}$	$C_{1,3}$	$C_{1,2}$	$C_{2,3}$	$C_{1,1}$	$C_{2,2}$	$C_{3,3}$	$C_{1,0}$	$C_{3,2}$	$C_{2,1}$	$C_{2,0}$	$C_{3,1}$	$C_{3,0}$

Table 3. INTRA AC, INTRA DC and 4×4 AC zigzag scan tables of HH band.

Index	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
INTRA AC	$C_{0,0}$	$C_{3,3}$	$C_{3,2}$	$C_{2,3}$	$C_{2,2}$	$C_{3,1}$	$C_{1,3}$	$C_{2,1}$	$C_{3,0}$	$C_{1,2}$	$C_{2,0}$	$C_{0,3}$	$C_{1,1}$	$C_{0,2}$	$C_{1,0}$	$C_{0,1}$
INTRA DC	$C_{0,0}$	$C_{3,3}$	$C_{3,2}$	$C_{3,0}$	$C_{3,1}$	$C_{2,3}$	$C_{2,2}$	$C_{2,1}$	$C_{2,0}$	$C_{1,3}$	$C_{1,2}$	$C_{1,0}$	$C_{1,1}$	$C_{0,1}$	$C_{0,3}$	$C_{0,2}$
4×4 AC	$C_{3,3}$	$C_{3,2}$	$C_{2,3}$	$C_{2,2}$	$C_{3,1}$	$C_{1,3}$	$C_{3,0}$	$C_{2,1}$	$C_{1,2}$	$C_{0,0}$	$C_{2,0}$	$C_{0,3}$	$C_{1,1}$	$C_{1,0}$	$C_{0,2}$	$C_{0,1}$

Comparing with the directional subsets proposed in intra prediction, these optimized Z scan tables show conformance in oriented properties obviously for each high band. By utilizing all of the proposed zigzag tables in the corresponding high band encoding, we get average 0.1 dB gain of PSNR both in intra and inter coding and 1.5% bit-rate reduction in comparison with the original scan results. The detailed results are shown in Table 4.

Table 4. Performance comparison of the proposed Z scan tables with the original.

Sequence	Intra		Inter	
	PSNR (dB)	Bitrate Reduction	PSNR (dB)	Bitrate Reduction
Bus	0.123	2.01%	0.1232	1.43%
Coastguard	0.0706	1.31%	0.1081	1.45%
Football	0.1039	1.52%	0.1478	1.71%
Foreman	0.0278	0.71%	0.0666	1.41%

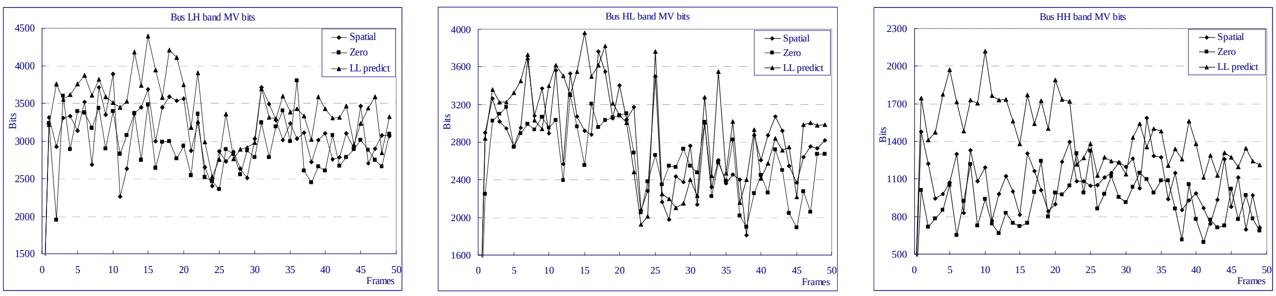
Furthermore, for some special applications of low complexity requirement, the number of proposed zigzag scan tables could be reduced to save memory cost both in encoder and decoder. As shown in the Table 1 to Table 3, the zigzag scan tables of INTRA AC are similar with those of INTRA DC in high frequency decompositions. Consequently, INTRA AC zigzag tables could be replaced by INTRA DC tables, which result in only 0.003 dB loss in PSNR, while 1/3 memory saving in scan tables.

3.3 Adaptive motion estimation

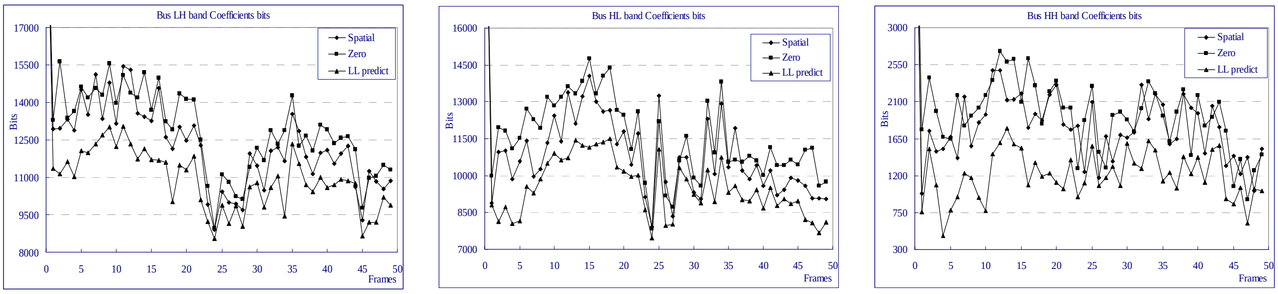
In H.264 motion estimation, the motion correlation among blocks is utilized by predicting current motion vector (MV) from its neighboring MVs. The final MV is achieved by refined motion search over reference frame and the difference between the final MV and its predictor is encoded. But in the overcomplete wavelet domain, such prediction always shows unsmoothed variation in motion field of high frequency bands. Though using LBS in forming reference frame for wavelet ME and MC could overcome the shift-variant property of wavelet transform, the MV prediction and MC in high frequency bands are still not so efficient as that in low frequency band.

In order to solve the problem, we divided the ME and MC problem of high frequency subbands into two aspects: one is the MV correlation among blocks inside frame, of which the MV predictor is the key point, which decides the start point of motion search and the coding bits spent on MV; the other is the relationship among the MVs of different subbands, especially the impact of the MV in LL band on high frequency subbands.

As mentioned before, though some kinds of spatial correlation exist in high frequency domain, the pixel distribution still contains irregular property especially in HH band. By analyzing the difference between the final MV of LL band and the three high frequency bands, three types of motion predictor are proposed to represent the different moving correlations in high bands. We named them as: the *Spatial* mode, using the neighboring MVs as predictor to represent the moving relevance inside each high frequency band; the *zero* mode, taking (0,0) as predictor to represent the irregular moving characteristics of high bands; and the *LL predict* mode, employing LL band MV as predictor to show the correlation between high band motion and low band. Comparing the performance of the three modes, we found that these three predictor types had their distinct advantages in saving MV cost or coefficients cost, but not in the both. The comparing results of Bus sequence are shown in Figure 4.



(a) (b) (c)



(d) (e) (f)

Figure 4: Motion vector bits and coefficients cost comparing results of Bus sequence coded in IPPP… GOP format for full sequence length skip two frames. Three motion predicting modes are compared in MV frame bits cost for: (a) LH band; (b) HL band; and (c) HH band; and in coefficients frame bits cost for: (d) LH band; (e) HL band; and (f) HH band.

Based on these statistical results, we propose to use the three types of motion predictor mode jointly and adaptively in the high frequency decompositions' ME to match the various moving properties and balance the cost in MV and coefficients simultaneously. The mode decision is determined by rate distortion optimization. Test results prove that using the proposed three MV predictor modes adaptively in the system contributes average 0.19dB gain in PSNR and 4.2% bit rate reduction. However, the ME complexity increases to three times in high bands comparing with the original and two bits overhead per partition are used to specify the predictor mode.

3.3.1 Adaptive motion prediction

By using the proposed three MV predictor modes, the ME in the high frequency bands are much more efficient than the original method. To further enhance the efficiency of the motion estimation, the correlation between the LL band and the high bands are utilized by proposing a new mode, *LL MV* mode. Table 5 clarifies the proposed *LL MV* mode together with the aforementioned three modes.

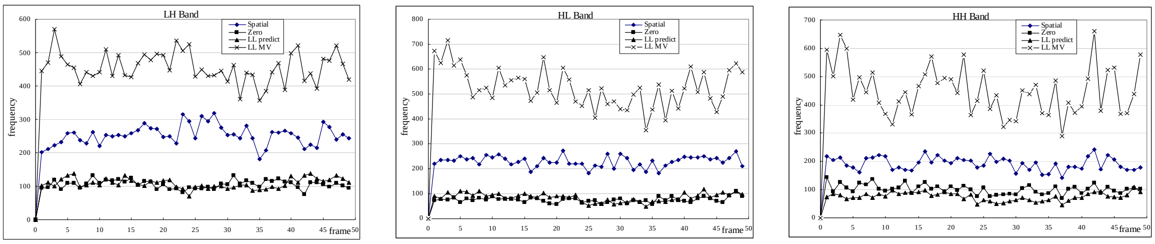
Table 5. Motion prediction mode for high frequency band.

Mode	Predictor	Refinement	Motion Cost	Final MV	Coded MV diff.
Spatial	Neighboring MV	Quarter pel refine	SAD+MVCost	Refined MV	Refinement – predictor
Zero	0,0				
LL predict	LL MV				
LL MV	No prediction	No refine	SAD	LL MV	0

The *LL MV* mode uses MV of LL band directly in high frequency subband for inter mode decision and MC. In the case of applying the *LL MV* mode individually to the proposed SSVC system, only LL MV field is generated for four subbands and no MV information needs to be encoded in high bands. Furthermore, the complexity of ME can be reduced to 1/4 since no motion search is necessary in high bands. Using four types of motion prediction mode adaptively in the high bands ME with rate distortion optimization, it achieves 0.35 dB gain in PSNR and 6.1% in bit rate reduction with the same system cost as three-mode adaptation.

3.3.2 Motion prediction optimization

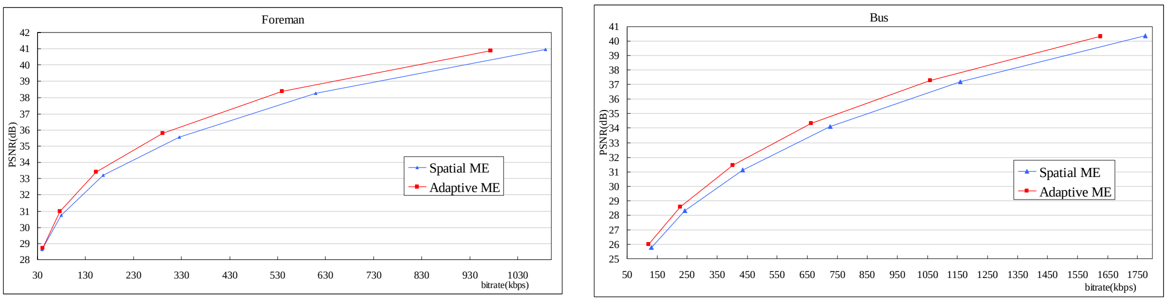
Though the adaptive motion prediction using the four proposed modes endures the system with attractive performance gain, the complexity increment makes it inappropriate to some resource-constrained applications. By further investigating the average usage of motion prediction modes in high frequency subbands as shown in Figure 5, we found that we could optimize the four-mode scheme into two-mode maintaining the moving correlation with close performance and less complexity.



(a) (b) (c)

Figure 5: Motion prediction mode analysis for high frequency band counting mode usage of spatial, zero, LL predict and LL MV mode for (a) LH band, (b) HL band, and (c) HH band.

It is noticeable that the *LL MV* mode and the *spatial* prediction mode for the three high bands ME occupy over eighty percent of total usage. Comparing the system performance in PSNR with four-mode adaptation, the two-mode scheme merely loss 0.06dB while with less mode payload. On the whole, the two-mode adaptive scheme contributes average 0.4dB gain in PSNR and 7.4% bit rate reduction, which is even better than the four-mode scheme. Furthermore, the complexity of ME is reduced to almost the same as the original. Figure 6 shows the comparing results between two-mode ME scheme and the original scheme of one spatial prediction mode.



(a) (b)

Figure 6: Performance comparison of proposed two-mode adaptive method with original spatial prediction method for: (a) Foreman, and (b) Bus.

4. EXPERIMENTAL RESULTS

Figure 7 shows the performance of the proposed SSVC schemes compared with H.264 and the original method. The test results presented below are performed on CIF “Bus” video sequence. One level 5/3 linear phase FIR filters are applied in the decomposition. The six-tap weiner filter is used in the interpolation in the proposed scheme. A single bit stream has been encoded at 30 frame/s (fps). The quantization parameters allocated to LH, HL and HH are three, four and five lager than LL’s.

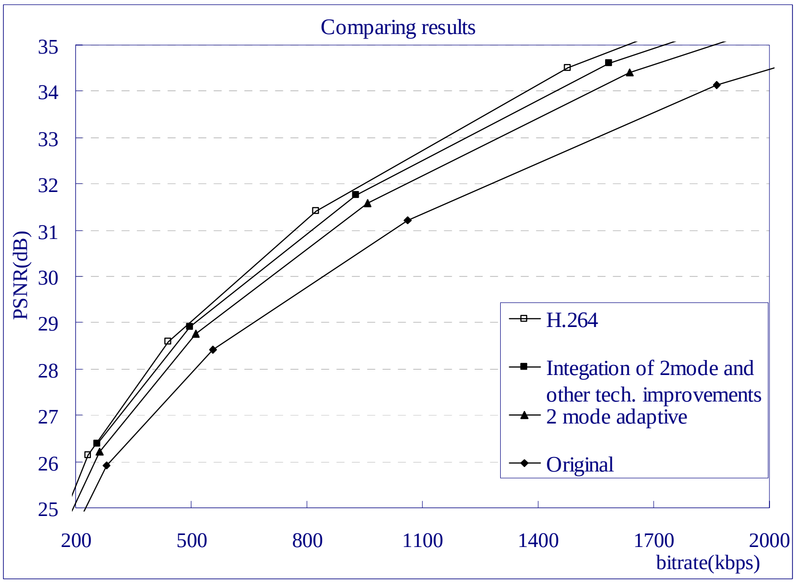


Figure 7: Performance comparison of several attempts.

As shown in the figure, four cases are compared. From the bottom to the top, the curves represent the original of one spatial prediction mode, the two-mode adaptive ME scheme, the integration of two-mode scheme with our proposed simplified intra prediction and zigzag scan tables, and the full resolution results encoded by H.264 respectively. Comparing with the H.264 results, our in-band spatially scalable video coding system achieves comparable performance. The performance results for different sequences testing the original and the system integrated all of the technique improvements by comparing with the H.264 are listed in Table 6, which prove the efficiency of technique improvements described in this paper.

Table 6. Performance results of proposed system with (named Proposed) and without (named Original) technique improvements comparing with H.264 full resolution encoded results.

Sequence	Original vs. H.264		Proposed vs. H.264	
	PSNR (dB)	Bitrate reduction	PSNR (dB)	Bitrate reduction
Bus	-1.2594	-32.50%	-0.197	-4.40%
Coastguard	0.4203	12.76%	0.4719	13.9959%
Foreman	-0.6888	-19.04%	-0.4579	-11.9774%
Football	-0.2043	-4.19%	0.0395	0.8238%
Average	-0.43305	-10.7427%	-0.03588	-0.39%

5. CONCLUSIONS

In this paper, a spatially scalable video coding system is proposed based on H.264 hybrid coding technique and in-band wavelet transform. The overcomplete discrete wavelet transform with low-band-shift method is used in the block-based wavelet domain in-band motion estimation and compensation. In order to further overcome the inefficiency of both motion predicting and coefficients encoding in high frequency subbands, the characteristics of directional relativity and motional correlation among blocks and subbands are analyzed, based on which the improvements in intra prediction, reordering scan and motion estimation are proposed. The optimized intra prediction using proposed direction subsets for each high band saves 30% bits for intra prediction mode. Three scan tables are presented for each high band quantized coefficient coding, which contribute 0.1 dB gain in PSNR to the system. The efficiency of motion estimation is greatly improved by four-mode adaptive scheme using the moving correlation between LL band and high bands together with the irregular moving properties inside the high bands. Considering the balance among system performance, algorithm complexity and memory cost, alternative technique improvements are additionally discussed. Integrating all of the technique improvements into our proposed SSVC system, the system shows over 0.4 dB gain in PSNR and 10.4% in rate reduction compared with the original scheme, which is valuable for scalable video coding and further investigation.

REFERENCES

1. Jens-Rainer Ohm, "Three-dimensional subband coding with motion compensation," IEEE Trans. Image Processing, 3, pp.559-571, 1994.
2. Seung-Jong Choi and John W. Woods, "Motion-compensated 3-D subband coding of video," IEEE Trans. Image Processing, 8, pp.155-167, 1999.
3. Marek Domanski, Adam Luczak, and Slawomir Mackowiak, "Spatio-temporal scalability for MPEG video coding," IEEE Trans. Circuits Syst. Video Technol., 10, pp. 1088-1093, 2000.
4. Y. Andreopoulos, A. Munteanu, G. Van der Auwera, P. Schelkens and J. Cornelis, "Scalable wavelet video-coding with in-band prediction-implementation and experimental results," in Proc. IEEE Int. Conf. Image Processing, pp. III-729-732, 2002.
5. Y. Andreopoulos, M. van der Schaar, A. Munteanu, etal, "Fully-scalable wavelet video coding using in-band motion compensated temporal filtering," in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 6-10 April, 3, pp. III - 417-20, 2003.
6. Claudia Mayer, "Motion compensation in-band prediction for wavelet-based spatially scalable video coding," in Proc. IEEE Int. Conf. Acoustics Speech Signal Processing, pp. III - 73-6, 2003.
7. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC), Mar. 2003.
8. Hyun-Wook Park and Hyung-Sum Kim, "Motion estimation using low-band-shift method for wavelet-based moving-picture coding," IEEE Trans. Image Processing., 9, pp. 577-587, 2000.
9. Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., 13, pp. 560-576, 2003.
10. J.R.Ohm, "Chapter on motion-compensated wavelet video", www.ece.ucdavis.edu/~mihaela/ohm_chapter.pdf.
11. Iain E.G. Richardson, H.264 and MPEG-4 video compression, John Wiley & Sons press, 2003.