Video Watermarking Algorithm for H.264 Scalable Video Coding

Jianfeng Lu，Li Li and Zhenhua Yang
Institute of computer Graphics and Image Processing, Hangzhou Dianzi University
Hangzhou, China
[e-mail: lili2008@hdu.edu.cn]
*Corresponding author: Li Li

Received October 10, 2012; revised December 11, 2012; accepted January 3, 2012; published January 29, 2013

Abstract

Because H.264/SVC can meet the needs of different networks and user terminals, it has become more and more popular. In this paper, we focus on the spatial resolution scalability of H.264/SVC and propose a blind video watermarking algorithm for the copyright protection of H.264/SVC coded video. The watermark embedding occurs before the H.264/SVC encoding, and only the original enhancement layer sequence is watermarked. However, because the watermark is embedded into the average matrix of each macro block, it can be detected in both the enhancement layer and base layer after downsampling, video encoding, and video decoding. The proposed algorithm is examined using JSVM, and experiment results show that is robust to H.264/SVC coding and has little influence on video quality.

Key words: H.264/SVC, video watermark, blind detect

The research work was supported by a research grant from Zhejiang province industrial projects for public service. (Grant No.2010C31006, 2011C21081)

http://dx.doi.org/10.3837/tiis.2013.01.004

1. Introductaion

H.264/SVC is an extension of H.264/AVC. It provides a mixed video stream with temporal scalability, spatial resolution scalability, and quality scalability, which makes it possible to extract video streams with different frame rates, frame sizes and frame qualities from a H.264/SVC coded video stream. Because of its scalability, H.264/SVC coded video has the ability to adapt to different network situations ranging from a cable network to wireless network, and user terminal situations ranging from a high-resolution desktop to low-resolution mobile phone. Thus, it is has become increasingly popular today.

Video watermarking can be defined as such a technique. It takes advantage of the temporal and spatial redundancy of video, embedding watermark bits into the video without affecting the video quality. When necessary, these watermark bits can be detected from the watermarked video as the basis for copyright protection or other uses. According to [1], video watermarking should meet the following triple basic conditions: imperceptibility, robustness, and capacity. Generally speaking, imperceptibility requires that watermark embedding will not cause noticeable distortion of the video, robustness requires that embedded watermarks can be detected correctly even after suffering various attacks, and capacity requires sufficient watermark bits allow copyright information to be embedded in the video.

In relation to H.264/SVC, video watermarking should take the scalability of the temporal and spatial resolution, as well as the quality, into consideration. These are briefly introduced.

1.1 Temporal Scalability

Temporal scalability is mainly related to the hierarchical B-pictures [2] [3], **Fig. 1** shows a hierarchical B-picture structure with a GOP size of 16. There are 17 frames in **Fig. 1**, frame 0 to frame 15 belongs to the current GOP, and frame 16 belongs to the next GOP. Each of the 16 frames in the current GOP is divided into 5 levels, marked level 0 to level 4. The core of temporal scalability involves dropping video frames belonging to a level one time, according to the frame rate of the extract bitstream. For example, when the frame rate is 30 fps (frames per second), the extract bitstream contains all 16 frames in the GOP, but when the frame rate is changed to 15 fps, the 8 frames belonging to level 4 will be dropped, and extracted bitstream will only contain the other 8 frames. As for video watermarking, temporal scalability requires that a watermark can be detected at all video frame rates. Thus, in order to ensure that watermarks can be detected correctly even under the minimum frame rate condition, watermarks should be embedded into those video frame that belong to level 0.

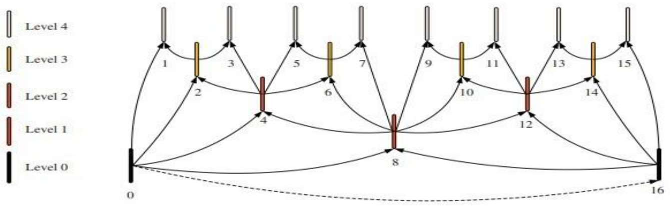


Fig. 1. Hierarchical B-picture structure with GOP size of 16[2]

1.2 Spatial Resolution Scalability

In H.264/SVC, spatial resolution scalability supports bitstreams with different spatial resolutions. In other words, we can extract both a CIF (352×288) resolution bitstream and a QCIF (176×144) resolution bitstream from a H.264/SVC coded bitstream if the enhancement layer resolution is CIF. With the currently specified profiles, the maximum number of enhancement layers in a bit stream is limited to 47, and at most two of these can represent spatial enhancement layers [4]. **Fig. 2** shows the simplified SVC encoder structure with tow spatial layer.

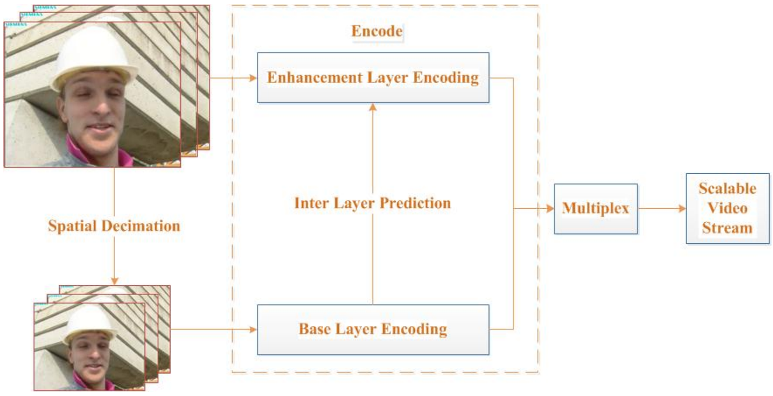


Fig. 2. Simplified SVC encoder structure with tow spatial layer

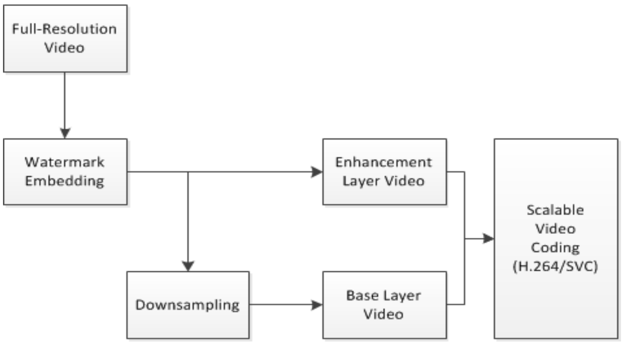
As you can see from **Fig. 2**, when there is spatial resolution scalablility in H.264/SVC, the video coding is carried out according to the following procedures. First, we perform spatial decimation on the enhancement layer of the original video sequence and get the base layer of the original video sequence. After that we can encode the base layer and enhancement layer. The base layer encoding is does what H.264/AVC does. When encoding the enhancement layer, inter-layer predictions are needed, which include interlayer intra-prediction, interlayer residual prediction, and interlayer motion prediction. By using these three inter-layer prediction methods, it can effectively reduce the redundancy between the spatial base resolution layer and enhancement resolution layer and improve the coding efficiency. As for video watermarking, spatial resolution scalability requires that a watermark can be detected in each spatial resolution layer.

1.3 Quality Scalability

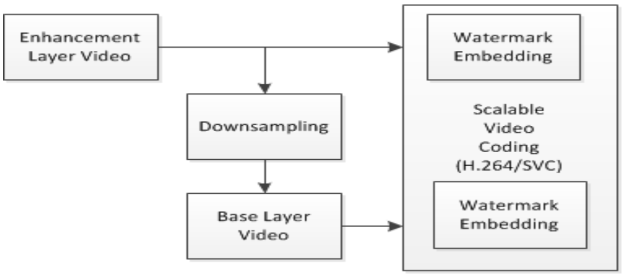
Quality scalability can be considered to be special case of spatial scalability with identical picture sizes for the base and enhancement layers [5]. There are two types of quality scalability supported by H.264/SVC: CGS (coarse-grain scalability) and MGS (medium-grain scalability). Because quality scalability is similar to using a different quantization step size to encode, if the video watermarking algorithm is robust enough to quantization, it will be robust to quality scalability. In this paper, we focus on the spatial resolution scalability of H.264/SVC and propose a blind video watermarking algorithm for the copyright protection of H.264/SVC coded video. In section 2, we briefly review the existing H.264/SVC video watermarking algorithm. In section 3, the proposed algorithm is presented in detail. Experimental results are presented in section 4, and our conclusions are given in section 5.

2. Video Watermarking For Resolution-Scalable H.264/SVC

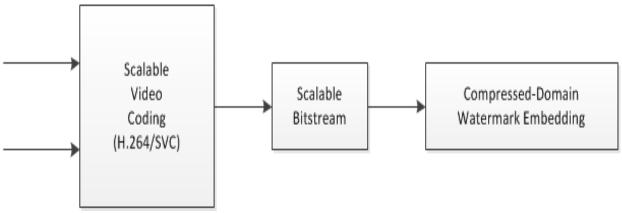
According to [6], the video watermarking algorithms for resolution-scalable H.264/SVC can be divided into three types: watermark embedding before video encoding, integrated watermark embedding and coding, and compressed domain embedding after encoding. **Fig. 3** shows the three different video watermarking types.



(a) Watermark embedding before encoding



(b) Integrated watermark embedding and coding



(c) Compressed-domain embedding after encoding

Fig. 3. Different embedding scenarios for resolution-scalable H.264/SVC video content [6][7]

In [6], Peter Meerwald and Andreas Uhl extended the robust H.264-integrated watermarking method in [8] to spatial scalability of H.264/SVC. They found that a watermark that is only embedded in the base resolution layer is not detectable in the enhance resolution layer. They solved this problem by adding an up-sampling watermark signal to the residual of the enhancement layer. The result showed that the watermark detection of their algorithm is reliable for both base and enhancement layers, and the watermark embedding reduced the video bitrate. Moreover, the watermark detection could not only be performed on the decoded video frames but could also be integrated into the H.264/SVC decoder.

Su-Wan Park and Sang-UK Shin proposed a combined encryption and watermarking scheme to provide the access rights and the authentication of the content in [9]. For the watermarking, they proposed a reversible watermarking scheme using the intra-prediction mode. Although the use of reversible watermarking did not cause visual quality degradation, their method has the bit overhead problem. In addition, the biggest drawback is their method can only be used in the base spatial resolution layer of H.264/SVC.

Robrecht Van Caenegem et al. gave attention to resilience against the spatial scaling applied in H.264/SVC and proposed a watermarking method [10]. They created a one-dimensional scaling invariant watermark host vector in the Foruier-Mellin domain [10], and embedded watermarks into the host vector. Experimental results show that their watermarking is robust against spatial scaling without video compression. When using H.264/SVC coding, the robustness of their method decreased, as a result of lossy compression.

Peter Meerwald and Andreas Uhl proposed a simple, frame-by-frame watermarking scheme as a vehicle for robustness experiments with scalability in [11]. In the proposed algorithm a two-level wavelet transforms with a 7/9 bi-orthogonal filter is decomposed on the luminance component of each video frame, and watermarks are separated and embedded in the approximation and each detail subband layer. For the coefficients $d_{l,o}(m,n)$ of the detail subband layer of each frame, an additive spread-spectrum watermark $w_l(n,m)$ is added as:

$$d_{l,o}^{'}(n,m) = d_{l,o}(n,m) + a \cdot s_{l,o}(n,m) \cdot w_l(n,m) \tag{1}$$ [11]

where a is the global strength factor and $s_{l,o}(n,m)$ is a perceptual shaping mask derived from a combined local noise and frequency sensitivity model [11]. l and o indicate the hierarchical level and orientation of the subband. Blind watermark detection can be performed independently for each hierarchical layer using normalized correlation coefficient detection. By applying a high-pass 3×3 Gaussian filter to the detail subbands before correlation, some of the host interference is suppressed, which improves the detection statics. Moreover, a different key is used to generate the watermark pattern for each frame [11].

Feng Shi, Shaohui Liu et al. [12] proposed a scalable and credible watermarking method. First, by considering the scalability and stability of the watermark, they proposed a unified method for selecting the embedding space of the SVC baseline. Then, they proposed an improved JND model in the wavelet domain. Finally, they proposed their adaptive watermark detection method.

3. Proposed Algorithm

The video watermarking algorithm proposed in this paper includes watermark embedding and watermark detection. The watermark embedding occurs before video coding, and the watermark detection occurs after video decoding.

3.1 Watermark Embedding

The structure of the watermark embedding is shown in **Fig. 4**. Watermarks are embedded in the luminance component of each video frame in the original enhancement layer of the SVC, and we can obtain the watermarked base layer of SVC by downsampling the enhancement layer. After this, we use the H.264/SVC encoder to code the enhancement layer and base layer and get the bitstream of H.264/SVC. The key point is the watermark algorithm should be robust to downsampling and video lossy compression.

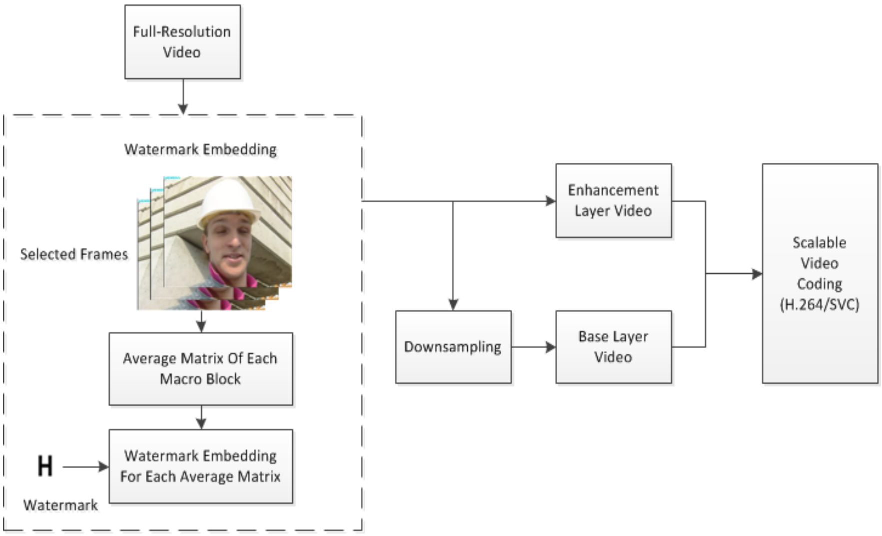


Fig. 4. Watermark embedding of the proposed algorithm

For the purpose of making the video watermark algorithm robust to downsampling, we construct the average matrix A (8×8) of each macro block MB (16×16).

1) Calculate the average matrix A:

$$A(i,j)=\frac{1}{4}\sum_{m=0,n=0}^{1}MB(2\cdot i+m, 2\cdot j+n), (0<=i,j<8) \tag{2}$$

2) Update the value of MB:

$$MB'(i,j)=MB(i,j)-A(\frac{i}{2},\frac{j}{2}), (0<=i,j<8) \tag{3}$$

Then, for the purpose of making the video watermark algorithm robust to video lossy compression, we use the following quantification method.

3) Transform the average matrix A into the DCT domain, and get A'.

4) Get sub matrix M (2×2) from A':

$$M(i,j)\in\{A'(i,j)\,|\,0<=i,j<=1\} \tag{4}$$

5) Perform SVD [13] decomposition on M, according to $M=USV^T$. Here, U and V are both 2×2, and S is a diagonal matrix that contains the singular values. We embed the watermark bit by the quantization of the biggest singular value $S(0,0)$.

6) Quantize $S(0,0)$ by quantization step λ, $S'(0,0)=S(0,0)/\lambda$.

7) Embed the watermark bits according to the following rule:

$$\begin{cases} S'(0,0)=S'(0,0)+1, if\,(watermark==0\,\&\&\,S(0,0)\%2!=0) \\ S'(0,0)=S'(0,0)+1, if\,(watermark==1\,\&\&\,S(0,0)\%2==0) \end{cases} \tag{5}$$

8) According to the above steps, restore the current macro block.

9) Repeat the embedding until all of watermarks have been embedded.

3.2 Watermark Detection

Before watermark detection, both the enhancement layer and base layer should be extracted from the watermarked H.264/SVC bitstream and decoded. Moreover, the watermark detection method for the enhancement layer and base layer are different. **Fig. 5** presents the watermark detection of the proposed algorithm.

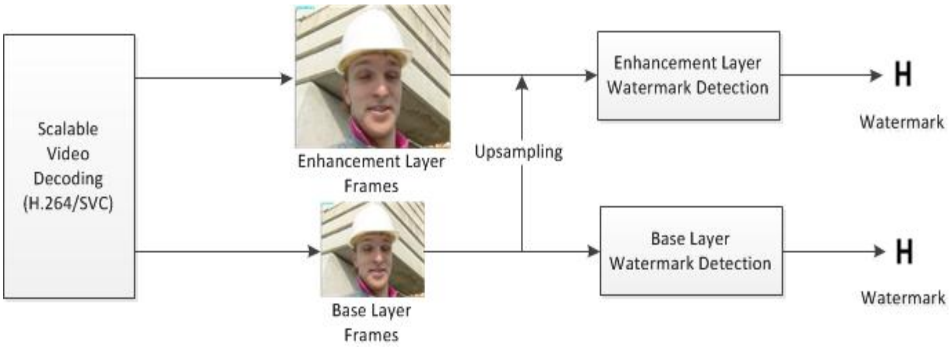


Fig. 5. Watermark detection of the proposed algorithm

3.2.1 Watermark Detection For Enhancement Layer

The watermark detection for the enhancement layer is similar to the watermark embedding process. The detection rule is as follow:

$$\begin{cases} watermark = 0, if\ (S^{'}(0,0)\%2 == 0) \\ watermark = 1, if\ (S^{'}(0,0)\%2 == 1) \end{cases} \tag{6}$$

3.2.2 Watermark Detection For Base Layer

There are two methods to detect watermarks in the base layer. The first method is upsampling the base layer into the spatial resolution of the enhancement layer, and then detecting the watermark as we did for the enhancement layer.

The second method is as follows:

1) Get a 8×8 block A from the luminance component of each base layer frame.
2) Transform A into the DCT domain, and get $A^{'}$.
3) Get sub matrix M (2×2) from $A^{'}$ as:

$$M(i,j) \in \{A^{'}(i,j) \mid 0 <= i, j <= 1\} \tag{7}$$

4) Perform SVD decomposition on M, according to $M = USV^{T}$. Here U and V are both 2×2, and S is a diagonal matrix that contains the singular values. We detect watermark bits by the quantization of the biggest singular value $S(0,0)$.

5) Quantize $S(0,0)$ with step size λ, $S^{'}(0,0) = S(0,0)/\lambda$.
6) Detect the watermark according to the following rules:

$$\begin{cases} watermark = 0, if\ (S^{'}(0,0)\%2 == 0) \\ watermark = 1, if\ (S^{'}(0,0)\%2 == 1) \end{cases} \tag{8}$$

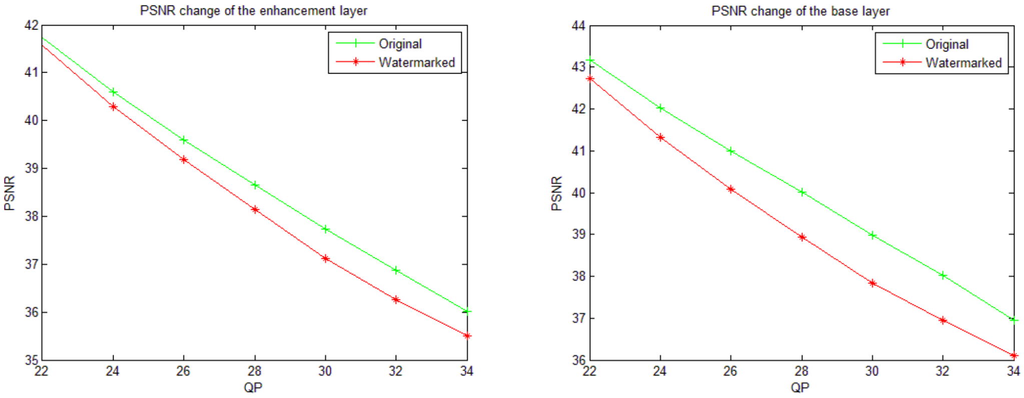
7) Repeat the detection until all of the watermarks have been detected.

4 Experimental Results

For the scalable watermark system, the key is the scalable property of the detection process [12]. That is to say, the watermark should be detected in both the base layer and enhancement layer of H.264/SVC spatial scalable coded video.

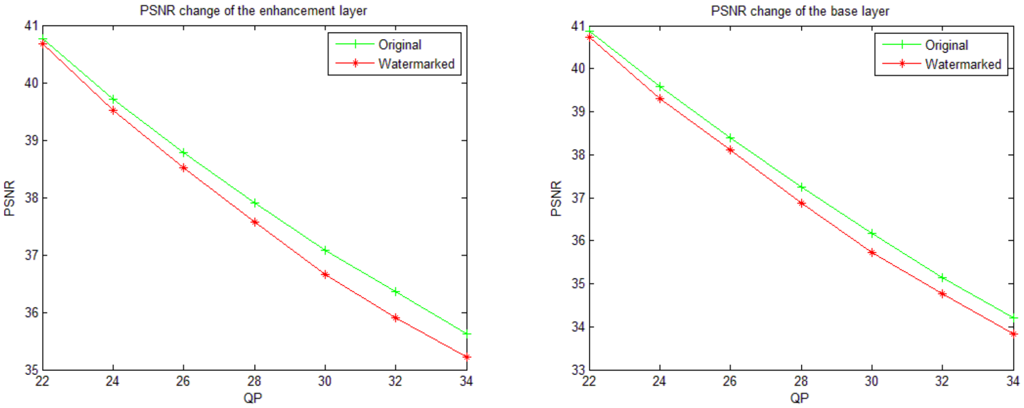
Here we test the proposed algorithm on JSVM_9_17, where the tested sequences are container_cif.yuv, foreman_cif.yuv, and news_cif.yuv. This means, for the spatial resolution, the enhancement layer is cif (352×288) and the base layer is qcif (176×144), and according to the section 3, we know that the watermark capacity for each video frame is 22 × 18 bits. When we tested the first 64 video frames of each sequence, the GOP size we set was 16, thus

1, 17, 33, 49 were I frames, and we only embedded watermarks in these 4 frames. And the quantization λ we used is 20. In all the experiments the QP factors for base layer and enhancement layer were equal, ranging from 22 to 34, with increments of 2 every time. **Fig. 6 - Fig. 8** shows the PSNR change after watermark embedding.



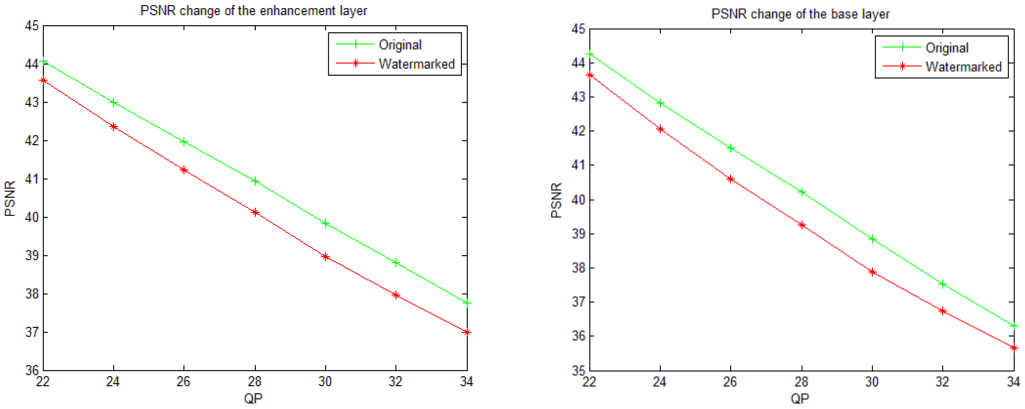
(a)Average psnr of the enhancement layer (b) Average psnr of the base layer

Fig. 6. Psnr of container



(a)Average psnr of the enhancement layer (b) Average psnr of the base layer

Fig. 7. Psnr of foreman



(a)Average psnr of the enhancement layer (b) Average psnr of the base layer

Fig. 8. Psnr of news

In **Fig. 6**, **Fig. 7**, **Fig. 8**, the green line (original) indicates the average psnr of the first 64 coded frames, when there was no watermark, and the red line (watermarked) indicates the average psnr of the first 64 coded frames when watermarked. When compared to the psnr values in [11], our reduction of psnr is stable, and is always in the range of 1 db.

Fig. 9, **Fig. 10**, **Fig. 11** present the bitrate change with the watermarking. The bitrate increase of our algorithm differs with the video sequence, spatial resolution and QP factor. For the container and news sequences, the bitrate increased more than foreman, and we found that this was because container and news are more flatter than foreman.

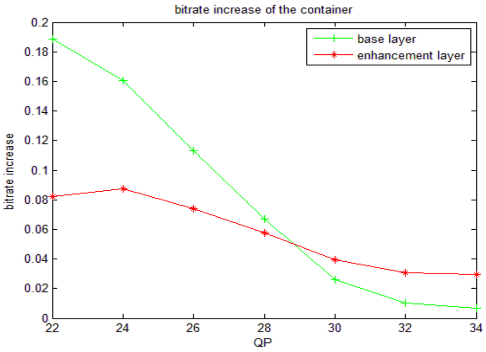


Fig. 9. Bitrate increase of the container

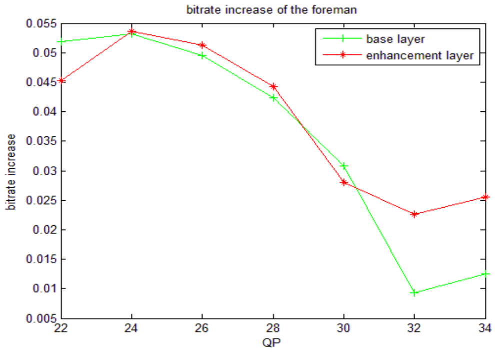


Fig. 10. Bitrate increase of the foreman

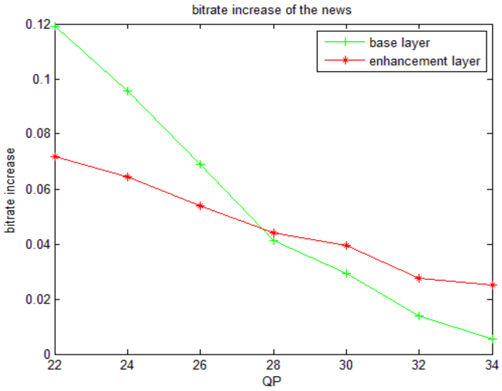


Fig. 11. Bitrate increase of the news

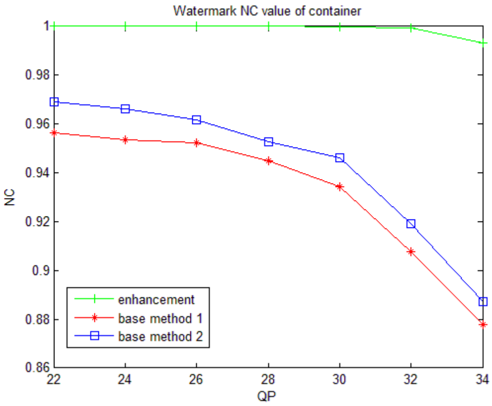


Fig. 12. Watermark NC value of container

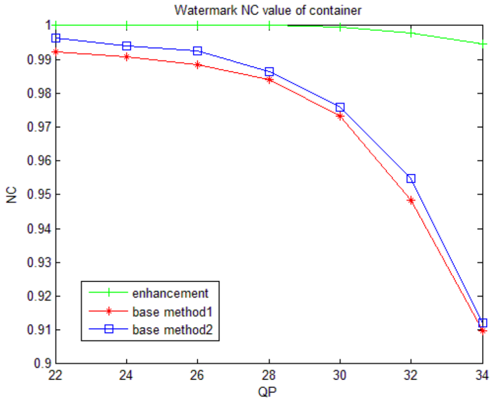


Fig. 13. Watermark NC value of foreman

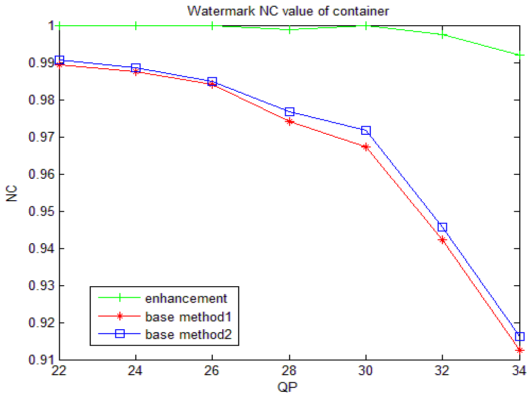


Fig. 14. Watermark NC value of news

Fig. 15. The compare of watermark images

Fig. 12, **Fig. 13**, **Fig. 14** show the average NC values of the watermarks detected from both the base and enhancement layers of the spatial resolution scalable H.264/SVC video stream. When QP is under 32, the watermark detection for the enhancement layer is almost perfect, and the detected watermark is nearly the same as the watermark we embedded. For the watermark detection of the base layer, the results of the two methods are almost the same, and when QP is under 32, the NC detection values are always above 0.9, which ensures that we can still identify the watermark with our eyes. **Fig. 15** presents the original watermark image on the left and a detected watermark image with an NC value of 0.88 on the right.

5. Conclusions

In this paper, we considered the spatial scalability of H.264/SVC, and designed a video watermarking algorithm. In the proposed algorithm, we only embed a watermark in the DCT domain of the average matrix of each macro block of the enhancement layer. After downsampling, video encoding, and video decoding, the watermark can be detected from both the enhancement layer and base layer. For some video sequences, the video bitrate increased slightly, but our algorithm is robust to H.264/SVC coding and has little influence on the video quality.

Reference

[1] Chung Yuk Ying, Xu Fang Fei and Faith Choy, "Development of Video Watermarking for MPEG2 video," In *Proc. of IEEE Region 10 Annual International Conference,* pp. 1-4, 2006. Article (CrossRef Link)

[2] Zhao Tiesong, Wang Hanli, Kwong Sam, Kuo C.-C. Jay and Halang Wolfgang A, "Hierarchical B-picture mode decision in H.264/SVC," *Journal of Visual Communication and Image Representation*, vol.22, no.7, pp.627-633, 2011. Article (CrossRef Link)

[3] Ohm, Jens-Rainer, "Advances in scalable video coding," In *Proc. of the IEEE*, vol.93, no.1, pp.42-55, 2005. Article (CrossRef Link)

[4] Schwarz H, "The Scalable Video Coding Extension of the H.264/AVC Standard," *Signal*

Processing Magazine, vol.25, no.2, pp.135-141, 2008. Article (CrossRef Link)

[5] Schwarz Heiko, Marpe Detlev, Wiegand Thomas, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.17, no.9, pp.1103-1120, 2007. Article (CrossRef Link)

[6] Peter Meerwald and Andreas Uhl, "An Efficient Robust watermarking Method Integrated in H.264/SVC," *Transactions on Data Hiding and Multimedia Security Vii Lecture Notes in Computer Science*, vol.71110/2012, pp.1-14, 2012. Article (CrossRef Link)

[7] Peter Meerwald and Andreas Uhl, "Robust watermarking of H.264/SVC-encoded video:quality and resolution scalability," *Digital Watermarking. Lecture Notes in Computer Science*, vol.6526 LNCS, pp.159-169, 2011. Article (CrossRef Link)

[8] M. Noorkami and R.M. Mersereau, "A framework for robust watermarking of H.264 encoded video with controllable detection performance," *IEEE Transactions on Information Forensics and Security*, vol.2, no.1, pp.14-23, 2007. Article (CrossRef Link)

[9] Park Su-Wan, Shin Sang-UK, "Combined scheme of encryption and watermarking in h.264/scalable video coding(SVC)", *1st International Symposium on Intelligent Interactive Multimedia Systems and Services*, vol.142, pp.351-361,2008. Article (CrossRef Link)

[10] Van Caenegem Robrecht, Dooms Ann, Barbarien Joeri, Schelkends Peter, "Design of an H.264/SVC resilient watermarking scheme," In *Proc. of SPIE-IS and T Electronic Imaging-Multimedia on Mobile Devices 2010*, vol.7542, 2010. Article (CrossRef Link)

[11] Peter Meerwald and Andreas Uhl, "Towards robust watermarking of scalable video," In *Proc. of SPIE – The International Society for Optical Engineering*, vol.6819, 2008. Article (CrossRef Link)

[12] Shi Feng, Liu Shaohui, Yao Hongxun, Liu Yan, Zhang Shenping, "Scalable and credible video watermarking towards scalable video coding," *11th Pacific Rim Conference on Multimedia, vol.6297 LNCS*, no.1, pp.697-708, 2010. Article (CrossRef Link)

[13] Mansouri A., Aznaveh Mahmoudi, Azar Torkamani, "SVD-based digital image watermarking using complex wavelet transform," *Sadhana-Academy Proceedings in Engineering Sciences*, vol.34, no.3, pp.393-406, 2009. Article (CrossRef Link)

Lu Jianfeng received his B.S. and M.S. degrees from the department of Mechanical Engineering in 1998 and 2001 respectively, and his Ph.D. degree in computer science department from the Zhejiang University in 2005. He works in the department of computer science in Hangzhou DIANZI University from 2005. His main research interests are in the areas of digital video and image processing and scientific visualization.

Prof. Li Li was born in 1973. Received computer PHD degree from ZheJiang University in 2004. Her research interests include computer image processingdigital water marking technology. She works in Hangzhou Dianzi University.

Yang Zhenhua was born in Hunan Province, in 1988. He received the B.S. degree from the Department of Computer Science, Hangzhou Dianzi University in 2010. He is currently pursuing the M.S. degree with department of Computer Science, Hangzhou Dianzi University too. His research interests include video watermarking and digital image process.