Towards robust watermarking of scalable video

Peter Meerwald and Andreas Uhl

Department of Computer Sciences, University of Salzburg
Jakob-Haringer-Str. 2, A-5020 Salzburg, Austria

ABSTRACT

This paper pulls together recent advances in scalable video coding and protection and investigates the impact on watermarking. After surveying the literature on the protection of scalable video via cryptographic and watermarking means, the robustness of a simple wavelet-based video watermarking scheme against combined bit stream adaptations performed on JSVM (the H.264/MPEG-4 AVC scalable video coding extension) and MC-EZBC scalable video bit streams is examined.

Keywords: Scalability, video watermarking, scalable bit stream, bit stream adaptation.

1. INTRODUCTION

With the advent of mobile devices capable of wireless transmission and ubiquitous presentation of multimedia content, scalable image and video coding is more and more employed to allow adaptation of a single multimedia stream to varying transmission and presentation characteristics. Each individual content consumer can extract the best video representation fitting his or her application from a single bit stream. A scalable video bit stream can be adapted to fit resolution, quality, and well as spatial or temporal presentation demands.

The JPEG2000 standard for image coding already addresses scalability by employing a wavelet transformation and embedded, rate-distortion optimal coding.[1] The previous JPEG standard had only limited scalability support (eg. progressive JPEG).

The current ITU and ISO video coding standards (H.264 and MPEG-4 AVC, resp.[2]) do not efficiently support rich scalability options, but work is under way to extended the standard with features to fully support scalable video coding (scalable video coding (SVC) extension to H.264/MPEG-4[3]). Furthermore, video coding based on motion-compensated temporal filtering (MC-TF) in combination with wavelet-based subband coding promises superior coding and scalability performance.[4]

Watermarking has been proposed to resolve copyright and content authentication and integrity questions for multimedia data by imperceptibly embedding information in the content.[5] Watermarks are designed to be detectable, even when the multimedia content is altered during transmission – an advantage over 'hard' cryptographic methods.[6]

In section 2, the MPEG-4 FGS standard,[7] H.264/MPEG-4 AVC scalability extension[3] and a 3D subband coding scheme (MC-EZBC[8]) are presented. In section 2.1 we briefly review video protection techniques based on cryptographic means which aim for scalability compliance.

Streaming and scalable multimedia transmission poses challenges as well as potentials for watermarking methods,[9] but has received little attention so far. Section 2.2 provides a literature survey with regards to watermarking methods explicitly supporting scalability.

We propose a simple watermarking scheme in section 3 which is used to evaluate the impact of video adaptations resulting from the before mentioned codecs. These results are presented and discussed in section 4 with concluding remarks in section 5.

Email: {pmeerw, uhl}@cosy.sbg.ac.at

2. SCALABLE VIDEO CODING

Scalable video codecs aim at encoding video in a single bit stream from which different representations, ideally in a rate-distortion optimal way, can be extracted. Among the scalability options are quality or signal-to-noise (SNR), resolution, spatial and temporal scalability.

There are several strategies to achieve scalability: layered coding, which is employed by MPEG-4 and its predecessors, embedded coding, used by 3D subband coders such as MC-EZBC, and hybrid methods, utilised by MPEG-4 Fine Granularity Scalability (FGS) and H.264/MPEG-4 SVC.

A single MPEG-4 bit steam can comprise two layers: a base layer and an enhancement layer which provides higher quality, higher resolution, and/or additional temporal frames. The aim of a scalable video codec is to optimise the video over a range of bit rates instead of a single bit rate. MPEG-4 FGS uses bit plane coding instead of run-length coding of residual data with the advantage to truncate the bit stream at any point to meet a bit rate budget.

H.264/MPEG-4 SVC efficiently supports multiple enhancement layers for combined scalability and improves motion-compensated predication by always resorting to the highest quality reference. However, there is still a 1 dB PSNR penalty when compared to non-scalable coding.[3]

The base layer can be coded in a H.264/MPEG-4 AVC compliant way, while the enhancement layers which add resolution, temporal or quality detail are coded with syntax according to the SVC proposal. For the remainder of this paper, we'll exclusively consider the Joint Scalable Verification Model (JSVM) investigating the scalability extensions based on H.264/MPEG-4 AVC.

All MPEG codecs adhere to the principle of hybrid codecs. In hybrid coding, one or more reference frames are used to compute a prediction of a given frame (involving motion and spatial estimation). The difference between the frame and its prediction is coded. The decoder's representation of the reference frame is used for prediction. The term 'closed-loop' system has been coined for this type of predictive coding. Since the encoder does not know whether the decoder has access to the enhancement layer information, efficiency suffers.

Within all the codecs of the MPEG family, temporal scalability is addressed by coding a hierarchical P or B frame structure.

An entirely difference approach to scalability is taken by 3D subband codecs[4] which were initially considered for JSVM but rejected because their architecture does not fit with MPEG's hybrid coding model. State-of-the-art 3D subband video codecs perform motion-compensated temporal filtering (MC-TF), followed by a spatial Wavelet decomposition. The hierarchical subband structure is then entropy coded, taking advantage of context modelling. 3D subband coding results in an embedded bit stream where video data is ordered according to perceptual significance. An embedded bit stream can simply be truncated at any point without degrading coding efficiency. The encoder does not have to be concerned about the state of the decoder for prediction ('open-loop' system). In this paper, we use the MC-EZBC codec as a representative of this class of codecs.

2.1 Scalable video protection

Video content is a undoubtedly a valuable digital commodity and its protection is essential to make commercial distribution viable and enable business scenarios such as super-distribution, video-on-demand, etc. Furthermore, content integrity verification and content authentication can add significant value. Encryption and cryptographic hashes have been proposed to meet these goals.

When encrypting a scalable video bit stream in a naïve way using a conventional cipher such as AES, the scalability properties of the bit stream are lost. Zhu et al.[10] surveys scalable encryption and authentication methods for JPEG2000 and MPEG-4 FGS employing cryptographic means. They note that when the content is adapted by intermediates in the distribution or transmission chain, these adaptations must not destroy the scalability property of the bit stream. However, it is desirable that no re-encryption or re-computation of signed hashes has to be performed when applying legitimate adaptations as these operations are computationally expensive and impose a key-sharing problem with potentially untrusted intermediates.

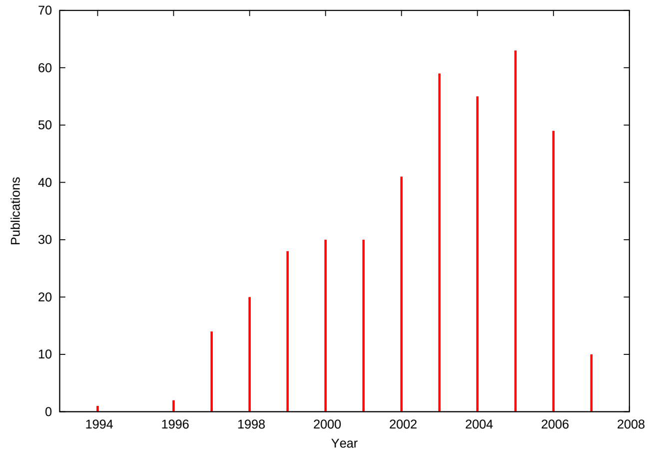


Figure 1. Number of publications on video watermarking per year. Source: author's bibliography.

Multiple video layers require multiple encryption keys and thus more complicated key management to support application scenarios. Eskicioglu et al.[11] proposes pre-positioned shared secrets and compares this approach against several other methods.

In a more recent work, Won et al.[12] examine access control and encryption for H.264/MPEG-4 SVC based on randomly inverting the video data's sign bit before entropy coding. Since scalability information is included in the network abstraction layer (NAL) encapsulating the video coding layer (VCL) data, bit stream adaptations can be performed by NAL unit dropping or cropping. Furthermore, a key management scheme reducing the number of keys for accessing the all the possible enhancement layers is proposed.

Scalable video authentication should accept all legitimate adaptations applicable to a scalable bit stream. This goal can be met by hard and soft- or content-based authentication techniques. The later methods build on perceptual hashing while the former employ a hierarchical signature structure, eg. Merkle hash tree.

Mukherjee et al.[13] propose a format independent encryption framework for scalable bit streams. Format independent bit stream adaptation is enabled by the use of MPEG-21 Part7 Digital Item Adaption (DIA) meta data augmenting the bit stream. Note that such a novel, format-agnostic approach is in contradiction to format-compliant technologies which might have a large installation base.

2.2 Video watermarking and scalability

Watermarking can provide a level of protection – a second line of defence – after the video has been decrypted. Therefore, it has been proposed to resolve content ownership claims or for traitor tracing in forensic fingerprinting applications. The later objective is achieved by embedding a fingerprint identifying the individual consumer. In addition, watermarks can be used to embed content integrity or authentication information for soft, i.e. according to a perceptual similarity measure, or hard verification. Another major advantage of robust watermarking over cryptographic means is its inherent tolerance to transmission errors.

Over the last decade, watermarking of multimedia content has received much attention. Video watermarking shares many aspects with image watermarking. However, peculiarities such as the human perception characteristics in the temporal domain, the high redundancy between frames giving rise to inter-frame collusion attacks, and the sheer volume of data to be processed – in real time for some applications – need to be considered separately. Doërr et al.[14] provide an overview of video watermarking techniques and challenges. More recent research has focused on the collusion attacks and countermeasures,[15,16] compressed-domain embedding,[17,18] with an emphasis on H.264/MPEG-4 AVC,[19–22] perceptual modelling of temporal human vision (HVS) characteristics[23,24] and signal-coherent embedding,[25,26] some approaches using 3D decompositions.[27,28]

Despite intense research in the area of video watermarking, see figure 1, watermarking in connection scalable video has neither received much attention nor is it well defined. In the following we identify six aspects of scalability and review related work in the respective directions.

Complexity scalability. As processing power increases at the detector's end due to technological innovation, more sophisticated algorithms or refined parameters can be used for watermark detection or synchronisation.[29] A straightforward example is exhaustive search to counter geometric attacks. In a similar sense can a watermark detector preempt its search if the watermark is found early in low distortion scenarios, thereby saving costly data transform operations. For image watermarking, Xia et al.[30] exploit the hierarchical nature of the wavelet decomposition to obtain watermark which is complexity and also inherently resolution scalable.

Detection progressiveness and robustness to quality scalability. We distinguish between quality scalability on the one, and resolution and temporal scalability on the other hand. The robustness of a watermark to coding at different bit rates, which is in most instances equivalent to a quantisation attack in the transform domain, is very well studied. All video codecs considered here form an embedded bit stream which can easily be truncated for scalability and an embedded robust watermark will thus support some degree of quality scalability. Resolution and temporal scalability pose more of a problem as spatial down-sampling and temporal adaptation pose also a synchronisation issue.

When watermark detection is integrated in the bit stream decoder, the watermark can be aligned with image components coded initially. Thus, watermark detection can commence early on while additional data is used to improve detection accuracy. A quality-progressive image watermark is proposed by Chen et al.[31] using the spectral selection mode of progressive JPEG images. Su et al.[32] integrate watermarking in the JPEG-2000 codec. Since the watermark is embedded in bit planes of significant coefficients which are transmitted first, it is possible to detect the watermark early, without decoding the entire image data. Lu et al.[33] claim that a watermark is scalable if it is detectable at low quality or resolution layers, i.e. progressive watermarking. Seo et al.[34] evaluate a scalable digital image watermarking scheme for protecting distance learning content and propose a progressive watermark embedded during wavelet-based image coding.

Watermark robustness to resolution and temporally scalable coding. An explicit notion of scalability first appears in the work of Piper et al.[35] They evaluate the robustness of different coefficient selection methods with regards to quality and resolution scalability in the context of the basic spread-spectrum scheme proposed by Cox et al.[36] Later, Piper et al.[37] combine resolution and quality scalability and argue that both goals can be achieved by exploiting the HVS appropriately.

Piper et al.[37] refine Lu et al.'s definition and states two properties for scalable watermarking along with numeric measures: detectability and graceful improvement. The detectability property states that a watermark shall be detectable in any version of the scaled content which is of acceptable quality. Graceful improvement refers to the desirable property that increased portions of scalable data shall be protected themselves as well as lead to more reliable watermark detection.

Pankajakshan et al.[16] discuss inter-frame collusion attacks by means of motion-compensated temporal filtering (MC-TF) and relates this attack to scalable video coding methods where the reduced frame rate video consists of low-pass temporal wavelet frames. The impact of this unintentional collusion attack is experimentally investigated for the MC-EZBC codec in section 4.

Scalable video authentication or integrity verification schemes should tolerate the distortion resulting from bit stream adaptations as legitimate manipulation. Sun et al.[38] presents a video authentication scheme based on DCT feature extraction combined with error correction coding (ECC) for scalable video streaming which takes into account re-quantisation, CIF-to-QCIF frame resizing and frame dropping.

Watermarking integrated with scalable coding. Alattar et al.[39] disclose a compressed domain watermarking scheme integrated with MPEG-4 which takes into account temporal and spatial scalability layers by watermarking them separately. They address the drift issue due to the prediction loop between spatial base and enhancement layer by subtracting the previously added watermark before adding a new one.

Wang et al.[40] propose a blind watermark embedded into MPEG-4 FGS bit planes for authentication of the enhancement layer. One bit is embedded by forcing the number of non-zero bits T_j per bit plane j and block to

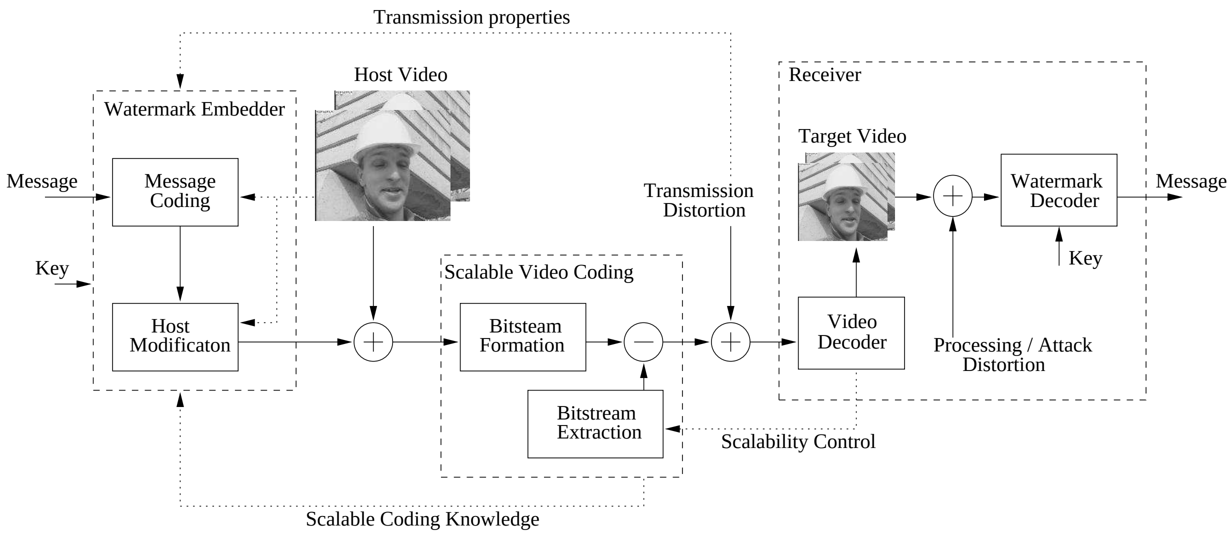


Figure 2. Watermark channel with scalable video coding.

even or odd depending on the watermark. If necessary, they choose to zero the bit yielding the longest zero run in order to meet the embedding condition, thereby improving coding efficiency. Unfortunately, the watermark does not depend on the image content and manipulation cannot be detected as long as $T_j \, mod \, 2$ is unchanged.

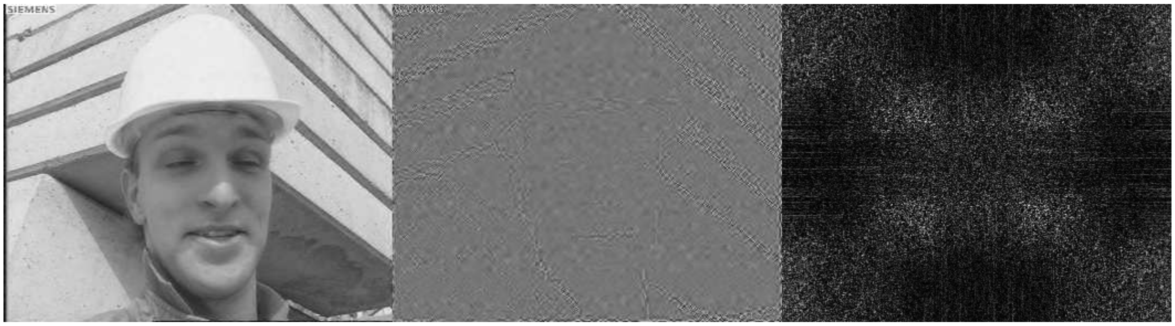
Distribution scalability. The fingerprinting scenario poses several challenges which become more severe as the number of users grows. In order to handle many users concurrently, fingerprinting can likely only be accomplished via compressed-domain watermarking and, to benefit from multi-cast distribution, the fingerprint would have to be embedded within the distribution network or at the receiver's end. Furthermore, fingerprint collusion resistance becomes more of an issue as the number of colluders as well as the number of users grows. He et al.[41] report detection results for 100 users applying an averaging or interleaving collusion attack out of a user set of 10 million for as little as 30 seconds of video data. Lin et al.[42] analyse how scalability layers can be exploited as side-information to improve the detection statistics in a forensic fingerprinting system.

Scalability and new application scenarios. Li et al.[43] propose a scalable audio watermarking scheme for the Advanced Audio Zip (AAZ) scalable audio codec part of MPEG-4. They embed a spread-spectrum watermark in both, the AAC base layer as well as the LLE enhancement layer. When the user does not have the key to the enhancement layer, only the watermarked and lower quality base can be decoded. The watermark distortion in either layer is compensated by the watermark in the opposite layer permitting lossless audio decoding when authorised to use the enhancement layer.

Chang et al.[44] combine encryption and watermarking to realize layered access control to a temporally scalable M-JPEG stream and consider an erroneous transmission channel. Watermarking is used to robustly embed the decryption key for the next frames. Four consecutive frames form a scalable group: the first frame constitutes the base layer, the remaining three frame the enhancement layer.

Lin et al.[9] discuss the challenges for watermarking associated with streaming and scalable video transmission. They raise the question where to embed the watermark: at the source, within the network, or at the receiver? Distortion due to transmission error or network loss is not perceptually bounded as typical attacks on the watermark. Error-concealment technique aiming at reconstructing the distorted video by estimating the damaged content can be seen as an attack. Ironically, data hiding has also been proposed to improve error resilience and aid in concealment.[45]

Most of the issues remain unaddressed in the literature and no results on the robustness of watermarking systems to scalable video codecs supporting combined scalability are available.



(a) watermarked (b) difference image (c) DFT difference

Figure 3. Embedding result on the first frame of the Foreman sequence: (a) watermarked, (b) difference image, (c) difference in DFT spectrum.

In the following sections, we focus on robust watermarking of scalable video. Figure 2 illustrates the watermark communication channel over a scalable video bit stream for blind watermarking. Assuming the availability of the original host video for watermark decoding is justified only for certain applications such as forensic fingerprint detection. The requirement to store the reference video is absurd or at least impractical for applications such as DRM, access control, and authentication due to security and data storage issues. Nevertheless, most watermarking algorithms reviewed above consider the non-blind case.

It is well known that by exploiting the host video as side-information at the encoder in message coding and watermark embedding, the negative impact of host signal noise on the watermark decoder performance can be cancelled.[5,46]

Using knowledge of the scalable video coding system or transmission channel may be beneficial for the watermarking system as well: number of supported resolution and temporal layers, using the same down-sampling filters to obtain low-resolution frames, denoising (MC-EZBC) and deblocking filters (JSVM), etc.

3. A SIMPLE SCALABLE WATERMARKING SCHEME

We propose a simple, frame-by-frame watermarking scheme as a vehicle for robustness experiments with scalable video coding. The luminance component of each frame is decomposed using a two-level wavelet transform with a 7/9 bi-orthogonal filter. Separate watermarks are embedded in the approximation and each detail subband layer. An additive spread-spectrum watermark $w_l(n,m)$ is added to the detail subband coefficients $d_{l,o}(n,m)$,

$$d'_{l,o}(n,m) = d_{l,o}(n,m) + \alpha \cdot s_{l,o}(n,m) \cdot w_l(n,m),$$

where α is global strength factor and $s_{l,o}(n,m)$ is a perceptual shaping mask derived from a combined local noise and frequency sensitivity model.[47] l and o indicate hierarchical level and orientation of the subband. Due to the high energy in the approximation subband, we adopt a spread-transform scalar Costa scheme (ST-SCS) for embedding. Blind watermark detection can be performed independently for each hierarchical layer using normalised 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 statistics.

A different key is used for each frame to generate the watermark pattern. More advanced key-schedule scheme should be employed in order to balance resistance against collusion versus watermark estimation and remodulation (WER) attacks.[16,25]

Figure 3 (a) shows the watermark embedding result on the first frame of the Foreman sequence. The difference image, (b), reveals the low-pass structure of the watermark, but also energy concentration in textured areas and around edges. In the DFT difference spectrum, (c), we observe the two high-frequency bands relating to the spread-spectrum watermark in the detail subbands as well as the low-frequency watermark in the approximation subband.

We will refer to above scheme as DWT-ST-SCS-N. It is designed to meet the following criteria for robust scalable watermarking: The watermark shall be detectable in the lowest resolution layer and reasonably low quality layers. Enhancement layer data is independently watermarked. Experimental robustness results are provided in the next section.

4. EXPERIMENTAL RESULTS

First, we embedded our DWT-ST-SCS-N watermark (see section 3) in the raw video data with an average strength of 36.4 dB PSNR. Perceptual shaping ensure the invisibility. Next, the watermarked video data in encoded in two scalable video bit streams using the H.264/MPEG-4 SVC reference implementation (JSVM version 9.1[*]) and the MC-EZBC[†] 3D subband codec. The JSVM bit stream has a GOP size of 16 and contains two resolution layer (QCIF and CIF at 30 frames/second) and three FGS layers. The quantisation parameter is set to 40. For MC-EZBC, the number of decomposition levels was set to 4.

Bit stream extraction was performed with the BitStreamExtract *<bit stream>* -e *<resolution>*@*<frame rate>*:*<bit rate>* and pull *<bit stream>* -s *<resolution layer>* -r *<bit rate>* commands, for JSVM and MC-EZBC, respectively. For JSVM, the bit steam was augmented with quality layer information using the QualityLevelAssigner -in *<bs>* -org 0 *<L0 video>* -org 1 *<L1 video>* -out *<bs>* -sei -mlql command.

The reported results relate to the first 32 frames of the Foreman sequence (CIF resolution, 352×288 pixel, YCbCr 4:2:0). Compression and rate allocation performance in terms of PSNR per frame is illustrated in figure 4 (a) for the MC-EZBC and JSVM bit stream adapted to 2000 kbit/second. Figures 4 (b) and (c) show the detector response for the watermark embedded in the approximation and highest detail subband, respectively. The PSNR per frame fluctuates wildly for both codecs and so does the detector response. The first frame per GOP has a considerable higher PSNR than the remaining frames.

Finally, in figure 4 (d) we experiment with MC-EZBC's denoising option and observe that denoising has a significant impact on watermark detection in the approximation subband (denoising is off per default). JSVM's default loop deblocking filter is only a minor influence and enabled per default.

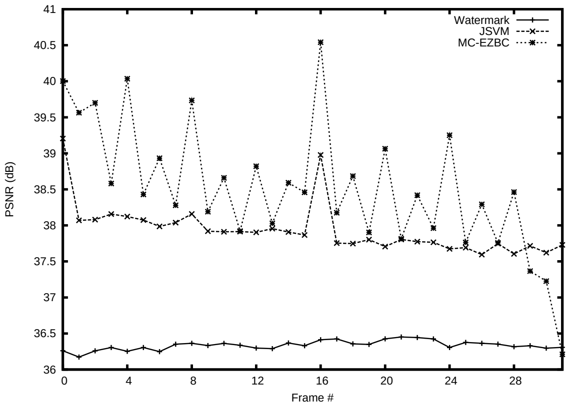
For the remaining experiments we only observe the approximation subband's watermark detector response without MC-EZBC denoising.

Robustness to quality scalability. Figure 5 (a) shows the watermark detector response averaged over the first 32 frames when extracting CIF resolution video and adapting the bit rate from 2500 to 250 kbit/second. The detector response decreases with the bit rate as expected. Higher values are obtained for MC-EZBC which is due to superior rate/distortion performance of MC-EZBC for this sequence (approximately 1.5 dB, compare with figure 4 (a) for results at 2000 kbit/second). The watermark from the JSVM bit stream can only be fully extracted for some frames, i.e. the first frame in a GOP and some P frames (compare with figure 4); on average the detector response drops to 0.8.

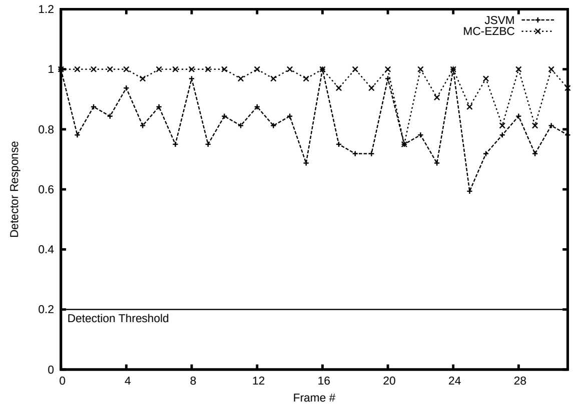
Robustness to resolution scalability. Figure 5 (b) provides the watermark detection results for the extracted QCIF sequence with bit rate adaptation from 1000 to 100 kbit/second. Note that detection from the JSVM bit stream outperforms MC-EZBC here. This is likely due to the direct watermark embedding in the low resolution video – the JSVM codec takes two separate video files as input for the two resolution layers whereas MC-EZBC derive the lower resolution by wavelet decomposition of the high resolution video.

[*]Available from http://ip.hhi.de/imagecom_G1/savce/index.htm.

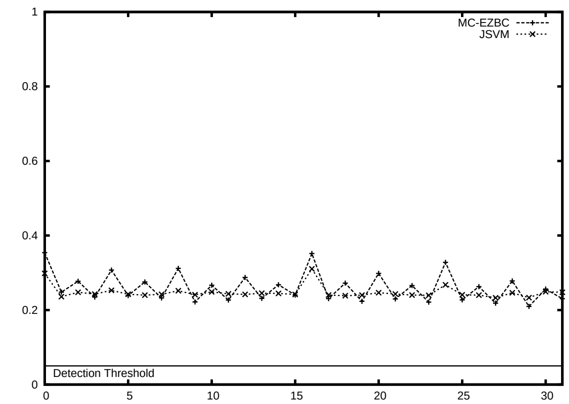
[†]ENH-MC-EZBC.zip, July 2005, available from http://www.cipr.rpi.edu/research/mcezbc/.



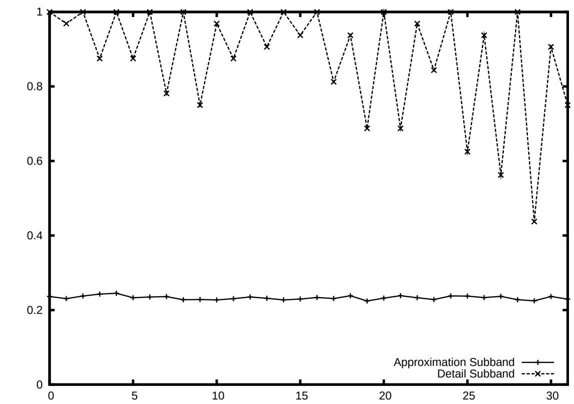
(a) PSNR



(b) Approximation Subband Watermark

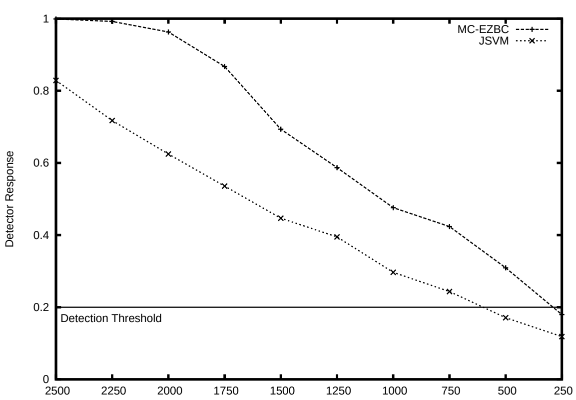


(c) Detail Subband Watermark



(d) MC-EZBC with Denoising

Figure 4. (a) JSVM and MC-EZBC PSNR at CIF resolution. Watermark detector response for the approximation (b) and highest detail subband (c) for the JSVM and MC-EZBC video sequence. (d) Watermark detector response when enabling the denoising option in MC-EZBC (off per default). All plots show the first 32 frames of the Foreman sequence, adapting the scalable bit stream to 2000 kbit/second.



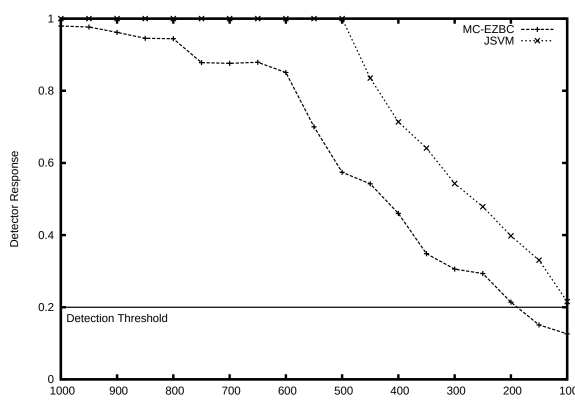
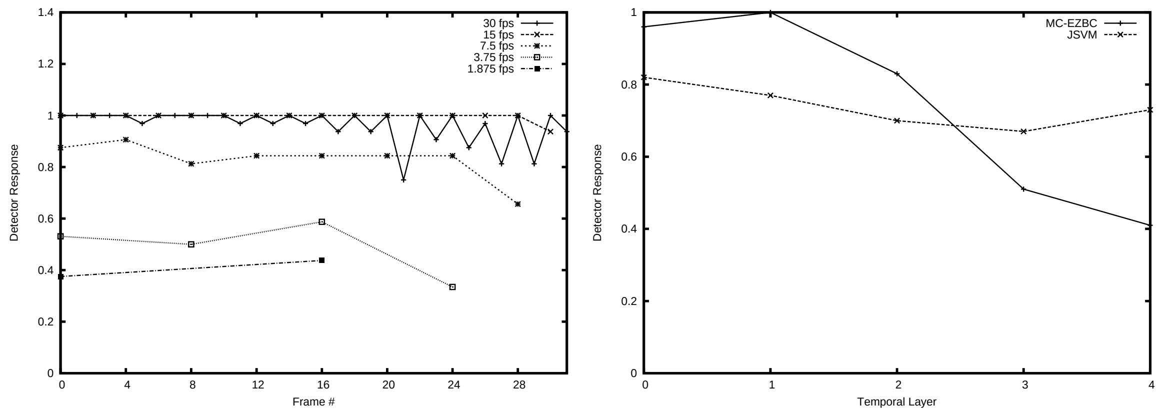


Figure 5. (a) Watermark detector response (averaged over the first 32 frames) when adapting the bit rate from 2500 to 250 kbit/second at CIF resolution, 30 frames/second. (b) Watermark detector response (averaged over the first 32 frames) after adapting the video to QCIF resolution for bit rates from 1000 to 100 kbit/second at 30 frames/second.



(a) MC-EZBC (b) MC-EZBC vs. JSVM

Figure 6. (a) Watermark detector response after temporal adaptation of the MC-EZBC bit stream to 30, 15, 7.5, 3.75, 1.875 frames/second, first 32 frames of the Foreman sequence (CIF). (b) Watermark detector response in temporal layers 0 to 4 (corresponding to 30 to 1.8.75 frames/second) for the MC-EZBC and JSVM bit steam, averaged over the respective frames in each temporal layer.

Robustness to temporal scalability. The detector response per frame obtained from the watermark in MC-EZBC's temporally adapted CIF video is shown in figure 6 (a). Interestingly, the detector response drops drastically for low temporal layers as predicted by Pankajakshan et al.[16] Adaptation from 30 to 1.875 frames/seconds have been performed. This is due MC-EZBC's temporal motion-compensated low-pass filtering which really is a motion-compensated temporal averaging (MC-TFA) attack.

Figure 6 (b) compares the watermark detector response obtained from the MC-EZBC and JSVM bit stream. The detector response has been averaged within each temporal layer – the JSVM's predictive coding does not hinder watermark detection.

5. CONCLUSION

We tried to refine and categorise the term 'scalable watermarkig' by reviewing the scalability aspect on watermarking in the literature. We have presented first watermark robustness results obtained with an experimental watermarking scheme designed for robustness scalability from a fully scalable H.264/MPEG-4 SVC and MC-EZBC bit stream. Scalable video coding broadens the range of unintentional, i.e. non-malicious attacks on the embedded watermark. The frame dropping attack often considered in video watermarking literature is insufficient to model temporal scalability due to temporal low-pass filtering. Motion-compensated temporal filtering/frame averaging must be considered an unintentional processing for 3D subband video codecs.

The proposed simple scalable watermarking scheme achieved robustness to combined adaptations of the investigated JSVM and MC-EZBC scalable bit streams. Further research is necessary in the area of watermarking integrated in scalable video codecs, esp. for the fingerprinting scenario. Scalability is a lot more than distortion on the watermark when considering applications of scalable video coding.

ACKNOWLEDGMENTS

This work has been supported by Austrian Science Fund project FWF-P19159-N13.

REFERENCES

1. D. Taubman and M. Marcellin, *JPEG2000 — Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, 2002.
2. I. E. G. Richardson, *H.264 and MPEG-4 video compression: video coding for next generation multimedia*, Wiley & Sons, 2003.

3. H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable H.264/MPEG4-AVC extension," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '06*, pp. 161–164, IEEE, (Atlanta, GA, USA), Oct. 2006.
4. J.-R. Ohm, "Advances in scalable video coding," *Proceedings of the IEEE* **93**(1), pp. 42–56, 2005.
5. I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*, Morgan Kaufmann, 2002.
6. A. Uhl and A. Pommer, *Image and Video Encryption. From Digital Rights Management to Secured Personal Communication*, vol. 15 of *Advances in Information Security*, Springer-Verlag, 2005.
7. W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Transactions on Circuits and Systems for Video Technology* **11**, pp. 301–317, Mar. 2001.
8. S.-T. Hsiang and J. W. Woods, "Embedded video coding using motion compensated 3-D subband/wavelet filter bank," *Signal Processing: Image Communication* **16**, pp. 705–724, 2001.
9. E. T. Lin, C. I. Podilchuk, T. Kalker, and E. J. Delp, "Streaming video and rate scalable compression: what are the challenges for watermarking?," *Journal of Electronic Imaging* **13**(1), pp. 198–208, 2004.
10. B. B. Zhu, M. D. Swanson, and S. Li, "Encryption and authentication for scalable multimedia: current state of the art and challenges," in *Internet Multimedia Management Systems V*, J. R. Smith, T. Zhang, and S. Panchanathan, eds., **5601**(1), pp. 157–170, SPIE, 2004.
11. A. Eskicioglu and E. J. Delp, "An integrated approach to encrypting scalable video," in *Proceedings of the IEEE International Conference on Multimedia and Expo, ICME '02*, **1**, pp. 573–576, (Laussanne, Switzerland), Aug. 2002.
12. Y. G. Won, T. M. Bae, and Y. M. Ro, "Scalable protection and access control in full scalable video coding," in *Proceedings on the 5th International Workshop on Digital Watermarking, IWDW '06*, *Lecture Notes in Computer Science* **4283**, pp. 407–421, Springer, (Korea), Nov. 2006.
13. D. Mukherjee, H. Wang, A. Said, and S. Liu, "Format independent encryption of generalized scalable bit-streams enabling arbitrary secure adaptations," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '05*, (Philadelphia), Mar. 2005.
14. G. Doerr and J.-L. Dugelay, "A guide tour of video watermarking," *Signal Processing: Image Communication* **18**, pp. 263–282, Apr. 2003.
15. K. Su, D. Kundur, and D. Hatzinakos, "Statistical invisibility for collusion-resistant digital video watermarking," *IEEE Transactions on Multimedia* **7**, pp. 43–51, Feb. 2005.
16. V. Pankajakshan and P. K. Bora, "Motion-compensated inter-frame collusion attack on video watermarking and a countermeasure," *IEE Proceedings on Information Security* **153**, pp. 61–73, June 2006.
17. B. G. Mobasseri, "Watermarking of MPEG-2 video in compressed domain using VLC mapping," in *International Multimedia Conference, Proceedings of the 7th Workshop on Multimedia and Security, MM-SEC '05*, pp. 91–94, ACM, (New York, NY, USA), Aug. 2005.
18. D. Simitopoulos, S. A. Tsaftaris, N. V. Boulgouris, A. Briassouli, and M. G. Strintzis, "Fast watermarking of MPEG-1/2 streams using compressed-domain perceptual embedding and a generalized correlation detector," *EURASIP Journal on Applied Signal Processing* **8**, pp. 1088–1106, July 2004.
19. G. Qiu, P. Marziliano, A. T. S. Ho, D. He, and Q. Sun, "A hybrid watermarking scheme for H.264/AVC video," in *Proceedings of the 17th International Conference on Pattern Recognition, ICPR '04*, pp. 865–868, IEEE, (Cambridge, UK), Aug. 2004.
20. M. Schlauweg, D. Pröfrock, T. Palfner, and E. Müller, "Quantization-based semi-fragile public-key watermarking for secure image authentication," in *Proceedings of SPIE, Mathematics of Data/Image Coding, Compression and Encryption VIII*, **5915**, pp. 5907–5916, SPIE, (San Diego, CA, USA), July 2005.
21. B. Mobasseri and Y. N. Raikar, "Authentication of H.264 streams by direct watermarking of CAVLC blocks," in *Proceedings of SPIE, Conference on Security, Steganography and Watermarking of Multimedia Contents IX*, **6505**, SPIE, (San Jose, CA, USA), Jan. 2007.
22. A. Golikeri, P. Nasiopoulos, and Z. J. Wang, "An improved scalar quantization-based digital video watermarking scheme for H.264/AVC," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS '06*, p. 5 pages, (Greece), May 2006.

23. A. Koz and A. A. Alatan, "Oblivious video watermarking using temporal sensitivity of HVS," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '05*, pp. 961–964, IEEE, (Genova, Italy), Sept. 2005.
24. M. Noorkami and R. M. Mersereau, "Perceptual video watermarking using motion estimation to exploit temporal masking," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '06*, pp. 1389–1392, IEEE, (Atlanta, GA, USA), Oct. 2006.
25. G. Doërr, J.-L. Dugelay, and D. Kirovski, "On the need for signal-coherent watermarks," *IEEE Transactions on Multimedia* **8**(5), pp. 896–904, 2006.
26. O. Harmanci and M. K. Mihcak, "Complexity regularized video watermarking via quantization of pseudo-random semi-global linear statistics," in *Proceedings of the 13th European Signal Processing Conference, EUSIPCO '05*, (Antalya, Turkey), Sept. 2005.
27. M. Carli, R. Mazzeo, and A. Neri, "Video watermarking in 3D DCT domain," in *Proceedings of the 14th European Signal Processing Conference, EUSIPCO '06*, (Firenze, Italy), Sept. 2006.
28. P. Campisi and A. Neri, "Video watermarking in the 3D-DWT domain using perceptual masking," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '05*, pp. 997–1000, IEEE, Sept. 2005.
29. I. J. Cox and M. L. Miller, "A review of watermarking and the importance of perceptual modeling," in *Proceedings of the 2nd SPIE Human Vision and Electronic Imaging Conference*, B. E. Rogowitz and T. N. Pappas, eds., **3016**, pp. 92–99, Feb. 1997.
30. X.-G. Xia, C. G. Boncelet, and G. R. Arce, "A multiresolution watermark for digital images," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'97)*, **1**, pp. 548–551, (Santa Barbara, California, USA), Oct. 1997.
31. T. P. chun Chen and T. Chen, "Progressive image watermarking," in *Proceedings of the IEEE International Conference on Multimedia and Expo, ICME '00*, pp. 1025–1028, (New York, NY, USA), July 2000.
32. P.-C. Su and C.-C. J. Kuo, "An integrated approach to image watermarking and JPEG-2000 compression," *Journal of VLSI Signal Processing* **27**, pp. 35–53, 2001.
33. W. Lu, R. Safavi-Naini, T. Uehara, and W. Li, "A scalable and oblivious digital watermarking for images," in *Proceedings of the International Conference on Signal Processing, ICSP '04*, pp. 2338–2341, IEEE, Aug. 2004.
34. J.-H. Seo and H.-B. Park, "Data protection of multimedia contents using scalable digital watermarking," in *Proceedings of the Fourth Annual ACIS International Conference on Information Science*, pp. 376–380, 2005.
35. A. Piper, R. Safavi-Naini, and A. Mertins, "Coefficient selection method for scalable spread spectrum watermarking," in *Proceedings of the 2nd Digital Watermarking Workshop*, *Lecture Notes in Computer Science* **2939**, pp. 235–246, Springer, Oct. 2003.
36. I. J. Cox, J. Kilian, T. Leighton, and T. G. Shamoon, "Secure spread spectrum watermarking for multimedia," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'97)*, **6**, pp. 1673–1687, (Santa Barbara, California, USA), Oct. 1997.
37. A. Piper, R. Safavi-Naini, and A. Mertins, "Resolution and quality scalable spread spectrum image watermarking," in *Proceeding of the 7th Workshop on Multimedia and Security, MMSEC '05*, ACM, (New York, NY, USA), Aug. 2005.
38. Q. Sun, H. Dajun, Z. Zhishou, and Q. Tian, "A robust and secure approach to scalable video authentication," in *IEEE International Conference on Multimedia & Expo, ICME '03*, **2**, pp. 209–212, (Baltimore, MD, USA), July 2003.
39. A. M. Alattar, E. T. Lin, and M. Celik, "Digital watermarking of low bit rate video," 2004. US Patent Application 20040125952, filed 08/06/2003.
40. C.-C. Wang, Y.-C. Lin, S.-C. Yi, and P.-Y. Chen, "Digital authentication and verification in MPEG-4 fine granular scalability video using bit-plane watermarking," in *Proceedings of the 2006 International Conference on Image Processing, Computer Vision, & Pattern Recognition, IPCV '06*, pp. 16–21, CSREA Press, (Las Vegas, NV, USA), June 2006.

41. S. He and M. Wu, "Collusion-resistant video fingerprinting for large user group," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '06*, pp. 2301–2304, IEEE, (Atlanta, GA, USA), Oct. 2006.
42. W. Lin, V. Zhao, and K. J. R. Liu, "Scalable multimedia fingerprinting forensics with side information," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '06*, IEEE, (Atlanta, GA, USA), Oct. 2006.
43. Z. Li, Q. Sun, and Y. Lian, "Design and analysis of a scalable watermarking scheme for the scalable audio coder," *IEEE Transactions on Signal Processing* **54**, pp. 3064–3077, Aug. 2006.
44. F.-C. Chang, H.-C. Huang, and H.-M. Hang, "Layered access control schemes on watermarked scalable media," in *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS '05*, pp. 4983–4986, IEEE, (Kobe, Japan), May 2005.
45. C. B. Adsumilli, M. C. Q. Farias, S. K. Mitras, and M. Carli, "A robust error concealment technique using data hiding for image and video transmission over lossy channels," *IEEE Transactions on Circuits and Systems for Video Technology* **16**, pp. 1394–1406, Nov. 2005.
46. I. J. Cox, M. L. Miller, and A. L. McKellips, "Watermarking as communications with side information," *Proceedings of the IEEE, Special Issue on Protection of Multimedia Content* **87**, pp. 1127–1141, July 1999.
47. M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Transactions on Image Processing* **10**, pp. 783–791, May 2001.