

Stereoscopic video watermarking: a comparative study

Afef Chammem · Mihai Mitrea · Françoise Prêteux

Received: 26 October 2012 / Accepted: 27 June 2013 / Published online: 20 July 2013
© Institut Mines-Télécom and Springer-Verlag France 2013

Abstract Despite the sound theoretical, methodological, and experimental background inherited from 2D video, the stereoscopic video watermarking imposed itself as an open research topic. Paving the way towards practical deployment of such copyright protection mechanisms, the present paper is structured as a comparative study on the main classes of 2D watermarking methods (spread spectrum, side information, hybrid) and on their related optimal stereoscopic insertion domains (view or disparity based). The performances are evaluated in terms of transparency, robustness, and computational cost. First, the watermarked content transparency is assessed by both subjective protocols (according to ITU-R BT 500-12 and BT 1438 recommendations) and objective quality measures (five metrics based on differences between pixels and on correlation). Secondly, the robustness is objectively expressed by means of the watermark detection bit error rate against several classes of attacks, such as linear and nonlinear filtering, compression, and geometric transformations. Thirdly, the computational cost is estimated for each processing step involved in the watermarking chain. All the quantitative results are obtained out of processing two corpora of stereoscopic visual content: (1) the 3DLive corpus, summing up about 2 h of 3D TV content captured by French professionals, and (2) the MPEG 3D video reference corpus, composed of 17 min provided by both academic communities and industrials. It was thus established that for a fixed

size of the mark, a hybrid watermark insertion performed into a new disparity map representation is the only solution jointly featuring imperceptibility (according to the subjective tests), robustness against the three classes of attacks, and nonprohibitive computational cost.

Keywords Robust stereoscopic watermarking · Spread spectrum · Side information · Hybrid watermarking · Stereoscopic disparity map · HD 3D TV

1 Introduction

Coined in the mid-1950s for audio signals, the stereoscopic world is nowadays the challenge in image/video processing. On the one hand, medical diagnosis, fault detection in manufactory industry, army, and arts, all of them consider multiview imaging as a key enabler for professional added value services. On the other hand, the worldwide success of movie releases (*Avatar*, *Alice in Wonderland*) and the deployment of 3D TV¹ chains made the nonprofessional user aware about a new type of multimedia entertainment experience.

However, the huge increase of distribution system for digital media leads to the increase of concerns over content copyright protection. Consequently, 2D content-inherited technologies such as cryptography, digital right management, watermarking, and network/access control are currently under reconsideration for their potential deployment for stereoscopic video protection.

The present paper takes the challenge of robust watermarking for stereoscopic video.

Watermarking can be considered as the most flexible property right protection technology, since it potentially supports all the requirements set by real-life applications

A. Chammem · M. Mitrea (✉)
Télécom SudParis, Institut Mines-Télécom, Évry, France
e-mail: Mihai.Mitrea@it-sudparis.eu

A. Chammem
e-mail: afef.chammem@telecom-sudparis.eu

F. Prêteux
MINES ParisTech, Institut Mines-Télécom, Paris, France
e-mail: francoise.preteux@mines-paristech.fr

¹ Throughout this paper, the terms stereoscopic video and 3D video are alternatively employed.

without imposing any constraints for a legitimate user. Actually, the watermarking protection mechanism, instead of restricting the media copy/distribution/consumption, provides means for tracking the source of the content illegitimate usage. In this respect, some extra information (a mark) is imperceptibly (transparently) and persistently (robustly) inserted in the original content.

From the application point of view, the watermarking challenge is to reach the trade-off between the properties of transparency, robustness, and data payload. A special attention is also devoted to the computational cost, mainly required by the real-time applications.

For 2D video-watermarking applications, several classes of insertion techniques already proved their efficiency. For instance, spread spectrum (SS) techniques [1, 2] are generally connected to excellent transparency properties while ensuring good robustness for a quite small size of the mark. Conversely, side information (SI) techniques [2–5] result in large sizes of the inserted marks while decreasing the transparency/robustness properties. Hybrid watermarking schemes, establishing synergies between the spread spectrum and side information principles, have also been reported [6]. This paper reconsiders these three classes of methods for stereoscopic video-watermarking purposes.

The stereoscopic video content also raises doubts about the optimal insertion domain. While the capturing and displaying of the 3D content is solely based on the two left/right views,² the bandwidth consumption reduction for their transmission/storage requires some alternative representations (e.g., disparity maps) to be designed. This issue is reflected in our paper at two levels. First, existing and new disparity map algorithms are benchmarked with respect to their watermarking potentiality. Secondly, the views and disparity maps are alternatively considered as insertion domains and the subsequent performances are compared. The differences between the view-based and disparity-based watermarking workflow are synoptically illustrated in Fig. 1.

The paper has the following structure: Section 2 outlines the different watermarking techniques considered in our study, while Section 3 stands for a state-of-the-art study of the stereoscopic video-watermarking approaches. Section 4 presents a new disparity map algorithm (further referred to as the 3D video new three-step search (3DV-NTSS)) and benchmarks it against the nowadays reference represented by the NTSS [7] and full-search MPEG (FS-MPEG) [8, 9]. Section 5 is devoted to the watermarking-related experimental results. Section 6 concludes the paper.

² The autostereoscopic displays and the depth image-based rendering (DIBR) imaging are out of the scope of the present paper.

2 Basic 2D watermarking insertion methods

A watermarking technique [2] consists in imperceptibly (transparently) and persistently (robustly) associating some extra information (a watermark) with some original content. From the theoretical point of view, a watermarking technique can be identified within the communication theory as a noisy channel. The mark is sampled from the information source and should be sent through a channel where a strong noise (the original content and the malicious attacks) acts. Under this framework, the properties of transparency and robustness are contradicting each other. A more transparent watermark can be achieved by decreasing its power at the transmitter, hence by reducing its robustness. The watermarking challenge consists in optimizing the transmission through this channel, i.e., in specifying a method achieving prescribed transparency and robustness while maximizing the size of the inserted mark.

This basic watermarking theoretical model yields to two main directions in watermarking: spread spectrum [1, 2] and side information [2–5]. The former considers the original content as a random noise source and maximizes the quantity of inserted data under joint constraints of noise (original content and malicious attacks) and inserted signal power. The latter considers the original content as a noise source known at the embedder [3] and maximizes the quantity of inserted data accordingly.

2.1 Spread spectrum

2.1.1 Principle

The SS methods have already been used in telecommunication applications (e.g., CDMA) and provided a good solution for very low-power signal transmission over noisy channels [1, 2]. Their principle consists in spreading the signal over a very large band (e.g., 100 to 10,000 times larger than the signal band), thus inserting a very low-power signal in any frequency sub-band.

The spread spectrum communications have two major benefits. First, the signal energy which is inserted into any frequency sub-band is very small, thus reducing the signal-to-noise ratio and the risk of perceptible artifacts. Secondly, the redundant watermark spreading over such a large band of frequencies provides robustness to many common signal distortions, as band-pass filter or addition of band-limited noise [1].

2.1.2 Embedding scheme

Among the large family of watermarking techniques, our study considers the method reported in [1], because of its well-recognized transparency and robustness properties. Let

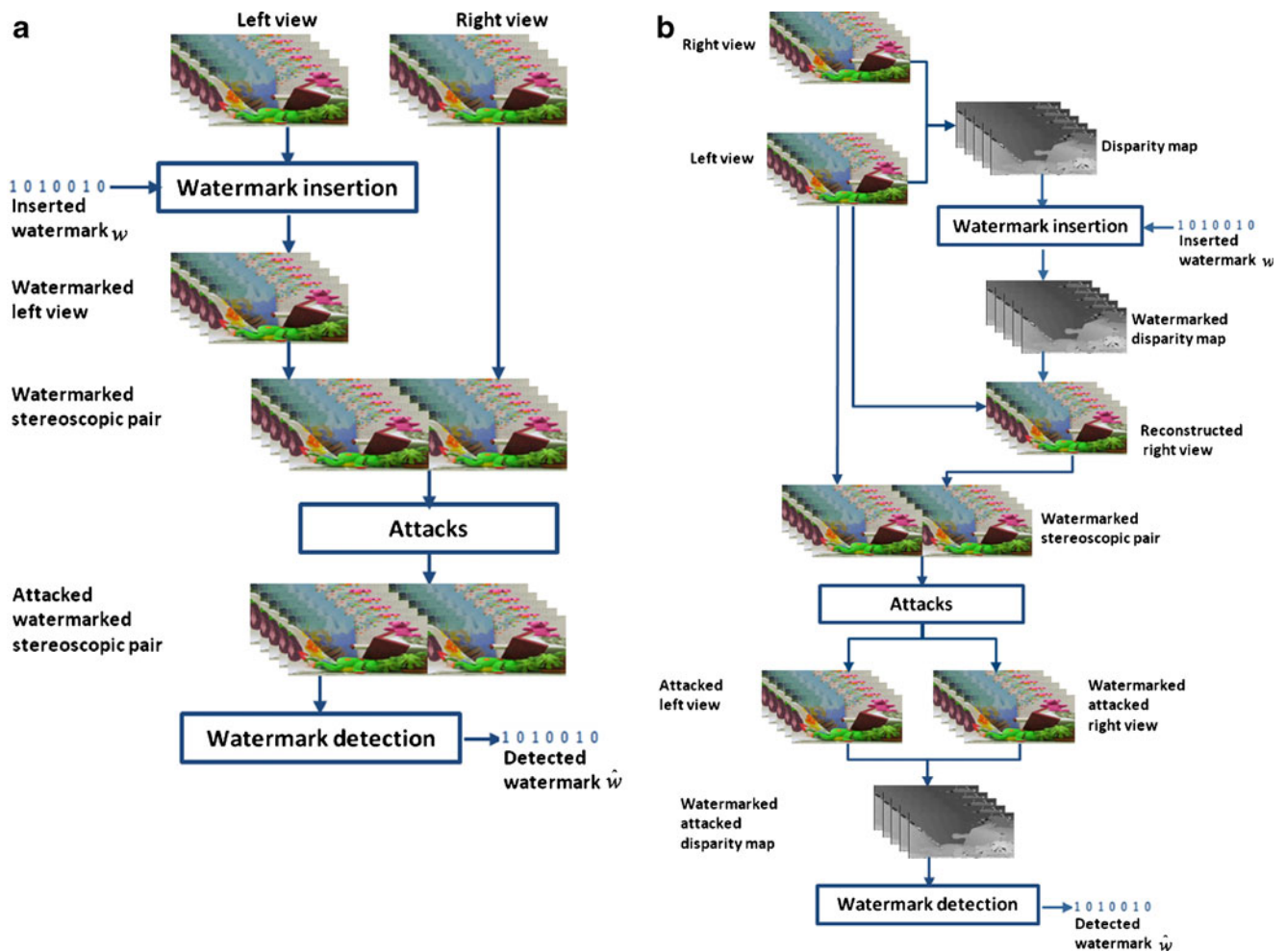


Fig. 1 View-based vs. disparity-based watermarking workflow

$M = \{m_i/m_i \in \{1,0\}\}$ be the additional information to be inserted in the original frame. Each bit m_i is spread by a large factor cr (the so-called chip rate), to obtain the spread sequence: $B = \{b_i\}$. The watermark w_i is added to image $C = \{c_i\}$, yielding a watermarked image $\hat{c}_i = c_i(1 + \alpha \cdot b_i)$, where α is the watermark strength.

2.2 Side information

2.2.1 Principle

Side information-based watermarking techniques take advantage of the fact that the original content is known at the embedder side (but unknown at the detector). This knowledge can be exploited at two levels. First, the *informed coding* methods encode the information to be inserted into the original content by a code word depending on that content [2, 5]. Secondly, the *informed embedding* methods modulate the watermark according to the original content [2–4]. In practice, informed coding and informed embedding can be individually or jointly deployed.

From the theoretical point of view, the side information watermarking is more sophisticated and should potentially outperform the blind coding and embedding methods. On the one hand, the informed coding would increase the transparency of the watermarking technique, and on the other hand, the informed embedding would grant robustness against a large range of attacks. However, the studies reported in the literature show that with respect to the spread spectrum method, the side information allows a significant increase of the data payload, generally at the expense of the robustness and/or transparency [2].

2.2.2 Quantization index modulation embedding scheme

The *quantization index modulation* (QIM) is an informed embedding watermarking method which proved its effectiveness for mono video-watermarking techniques, by achieving good robustness while keeping within practical limits the perceptual distortions [10, 11]. The method starts by modulating an index or a sequence of indexes with the message m to be embedded and then quantizes the host data

c_0 by using the associated quantizer or sequence of quantizers.

Initially designed under the binary framework, the QIM methods were generalized to *multisymbol QIM* in [10]. Be there a binary message to be inserted, instead of directly inserting it, a message m encoded into an s -ary alphabet $D = \{-(s-1)/2, -(s-2)/2, \dots, 0, \dots, (s-2)/2, (s-1)/2\}$ is considered so as to increase the data payload by a factor of $\log_2(s)$.

For a host signal c_0 and a message m , the watermarked signal sample y is computed by

$$y = c_0 + \alpha \cdot q$$

$$\text{where } w = \alpha \cdot q \text{ and } q = Q_{\Delta} \left(c_0 - \Delta \left(\frac{m}{s} + k \right) \right) - \left(c_0 - \Delta \left(\frac{m}{s} + k \right) \right)$$

where Δ is a fixed quantization step size, k is a random key, and α is a fixed parameter, $0 < \alpha \leq 1$. The standard quantization operation is defined as

$$Q_{\Delta}(c_0, \Delta) = \Delta \text{Round}(c_0/\Delta).$$

At the decoder, the embedded message bit is recovered by a scalar quantization of the received signal sample r (a corrupted version of y).

The $Y(m)$ detection variable is computed as follows:

$$Y(m) = Q_{\Delta}(r - k\Delta) - r + k\Delta$$

The decision rule is given by

$$\begin{cases} -I_{\text{sup}}(m) \leq Y(m) \leq -I_{\text{inf}}(m) \\ \alpha > \frac{s-1}{s} \end{cases}$$

$$\text{where } I_{\text{sup}}(m) = \frac{\Delta((1-\alpha)s+2m)}{2s} \text{ and } I_{\text{inf}}(m) = \frac{\Delta((\alpha-1)s+2m)}{2s}.$$

2.3 Hybrid watermarking

2.3.1 Principle

The hybrid watermarking methods [2, 6] combine spread spectrum and side information concepts, thus trying to benefit from both the robustness and transparency of the spread spectrum methods and the increased data payload of the side information methods. In this respect, the present study reconsiders the *IProtect* method [12, 13].

2.3.2 *IProtect* embedding

IProtect considers the watermarking procedure as an optimization problem, where the robustness is maximized under transparency and data payload constraints.

IProtect starts by setting the minimal and maximal acceptable level of the targeted robustness, denoted by R and R_r , respectively. Then, the embedding procedure iteratively searches for the targeted robustness (see Fig. 1) while

keeping the transparency superior to a given limit (e.g., peak signal-to-noise ratio (PSNR) >31 dB). The method convergence is heuristically obtained by guiding the searching procedure according to the statistical behavior of the would-be attacks.

The embedded message m (whose size gives the data payload) is encoded by means of a modified trellis code and is then inserted into the initial host content c_0 (Fig. 2). The issue is to find a c_w watermarked vector which is as close as possible to the c_0 vector (thus ensuring transparency) and for which the Viterbi decoder produces the same output as for the m vector, regardless of the type of the applied attack (thus ensuring robustness). The c_w vector is computed by an iterative algorithm. At the first iteration, $c_w = c_0$. Further on, for each iteration, a vector denoted by b is computed by applying the Viterbi decoder to $c_w + n$ and by encoding the resulting bits by a modified trellis method [2, 13, 14]. Here, n is a vector sampled from a noise source modeling the channel perturbations (i.e., the attacks); within *IProtect*, the n vector is a combination of a Gaussian distributed n_g vector (the default assumption in watermarking) and of an n_a vector whose components are generated so as to follow the probability density function modeling the combination of geometric attacks [15]. As the convergence of such an algorithm cannot be theoretically ensured and is subject to statistical evaluations, the iteration number should have some a priori upper limits (denoted in Fig. 2 by N_i and N_j).

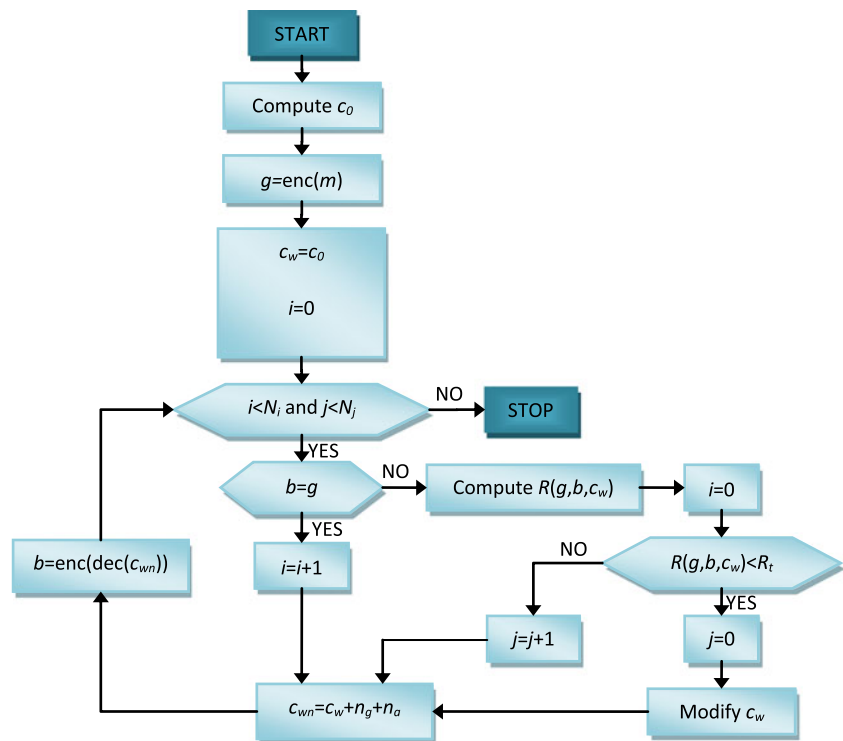
3 Stereoscopic video watermarking: state of the art

The 3D video state of the art brings to light several studies addressing the visual stereoscopic content watermarking. While the insertion method itself is always directly inherited from still image/mono video (being either of spread spectrum or of informed embedding type), these studies can be structured into *view-based* and *disparity-based* methods, according to the domain in which the mark is actually inserted.

The predilection direction in the literature is represented by the view-based watermarking approaches, which are currently deployed for stereoscopic still images [16–21].

For instance, in their incremental studies [16, 17], Dong-Choon et al. address several issues connected to the spread spectrum stereo image watermarking. In [16], a mark of 1,024 bits (representing a visual logo) was embedded in the discrete cosine transform (DCT) domain of each of the right views of a stereo sequence of 25 image pairs with a resolution of 256×256 pixels. Both transparency and robustness performances have been analyzed in terms of PSNR. Thus, the watermarked views feature an average PSNR of 34.89 dB when compared to their corresponding original views. While no information is provided about the actual applied attacks,

Fig. 2 IProtect embedding. The $\text{enc}(\cdot)$ and $\text{dec}(\cdot)$ functions denote the trellis encoder and the Viterbi decoder, respectively. The n_g and n_a terms represent the Gaussian and non-Gaussian noise components, respectively, while m denotes the inserted message



the robustness is evaluated as the PSNR between the recovered and the inserted logos and an average value of 19.04 dB was obtained in this respect. In [17], the experiment was resumed in the discrete wavelet transform (DWT) domain. This time, a sequence of three image pairs of 512×512 pixels was considered as original content and a logo of 64×64 pixels as watermark. While the average transparency was increased by 3.92 dB (reaching now the limit of 38.81 dB), the robustness was decreased by 2.36 dB (i.e., lowered at 16.68 dB).

Kumar et al. [18] also considered a spread spectrum-based watermarking method. The disparity map is first estimated in the DWT domain and then embedded as a mark in the left view. The embedding procedure is optimized by a genetic algorithm and takes place in the singular value decomposition (SVD) of the left image DWT. The experiments have been performed on five gray-level stereo images with a resolution of 512×512 pixels. The obtained transparency was expressed by a PSNR of 42.88 dB. The robustness against average filtering, rotations, resizing, and Gaussian noise addition was assessed by computing the normalized cross correlation (NCC) value between the original and extracted watermarks, namely, 0.91, 0.91, 0.90, and 0.89, respectively.

In the study by Bhatnager et al. [19], the watermark is represented by the disparity map, computed this time directly in the pixels domain. This disparity map is further inserted by a SS technique applied in the SVD of the left image discrete fractional Fourier transform (FrFT). The experimental validation considered a corpus of three stereo images (two image pairs of 256×256 pixels and one image pair of 512×512 pixels). The

transparency evaluation shows an average PSNR of 45.92 dB. This excellent value is obtained at the expense of the robustness which is now reduced at NCC values between the original and the extracted watermark of 0.64, 0.71, 0.69, 0.98, 0.57, 0.63, and 0.46 for the average filtering, median filtering, resizing, JPEG compression, cropping, rotation, and additive noise attacks, respectively.

Campisi [20] advances a semi-fragile stereo watermarking approach based on the QIM insertion method performed in the DWT domain. The watermark payload is 2,000 bits per frame. Experiments show that the advanced method is robust towards JPEG and JPEG2000 compression and fragile with respect to other signal manipulations. The fragility property was assessed by computing the values of bit error rate (BER); the following numerical values are obtained: 0.07 after Gaussian filtering, 0.11 after median filtering, 0.38 after row/column removal, 0.30 after a 0.25° rotation, and 0.39 after 0.50° rotation. No information concerning the transparency and the experimental data base is provided.

The study advanced by Yu et al. [21] embeds the watermark into both the left and the right views of the stereo pair, by taking into account some intra and inter-block statistical relationships established by combining the DWT and DCT. A parity quantization is also designed for handling the cases in which such a relationship does not hold. During the experiments, a binary logo of 64×64 pixels is embedded in a stereo frame of 640×480 pixels. The transparency is evaluated by the PSNR value between the original and the watermarked images; values of 52.14 and 51.99 dB are

obtained for the left and the right views, respectively. The robustness was evaluated in terms of the watermark recovering rates (WRR) (referred to as HC in [21]) given by $WRR = 1 - \sum w \oplus w' / m \times m$, where \oplus is exclusive OR, w denotes the original binary watermark of $m \times m$ size, and w' denotes the recovered watermark. WRR values of 0.94, 0.90, 0.94, and 0.81 are obtained after applying a JPEG compression, a salt and pepper noise, a median filtering, and a cropping, respectively. Note that $WRR = 1 - BER$.

The disparity-based stereoscopic watermarking schemes can be represented by the study reported in [22]. The insertion technique combines spread spectrum principles and low-density parity check error-correcting codes. The experiments are performed on four video sequences: three of them are composed of 22 stereo frames of 512×512 pixels while the fourth one has 192 frames of 720×576 pixels. The transparency is assessed by computing the average image degradation index $\bar{\delta} = 0.06$. This index is defined by $\bar{\delta} = \bar{\alpha} - \bar{\beta}$, where $\bar{\alpha}$ and $\bar{\beta}$ denote the average PSNR of the nonwatermarked and watermarked video images, respectively (no details are provided on how the nonwatermarked sequence average PSNR is computed). The robustness is assessed in terms of the watermark detection ratio (WDR), which was evaluated for each video and defined by $WDR = e_0 e_1 / m_0 m_1$, where m_0 and m_1 are the numbers of “0” and “1” of the original watermark, respectively, while e_0 and e_1 are the numbers of the extracted “0” and “1” from the recovered watermark, respectively. The reported WDR values are larger than 0.78 after recoding and 0.74 after both recoding and noise addition.

Table 1 presents a general view on these various approaches. When considering them for a real-life application (e.g., HD 3D TV real-time content protection), the following main limitations can be identified:

- The structure and the size of the processed corpora are too small to ensure generality and statistical relevance for the results;
- The selection of the insertion domain was rather the experimenter’s choice than the result of an objective study;
- The transparency evaluation was solely guided by the PSNR values; no other objective transparency method nor subjective evaluation is reported;
- The robustness against geometric attacks is not yet achieved;
- The computational cost was never investigated.

The present study addresses these above-mentioned issues:

- All the results consider two corpora, further referred to as 3DLive [23] (<http://3dlive-project.com>, summing up about 2 h of 3D HD TV content) and MPEG [24] (<http://sp.cs.tut.fi/mobile3dtv/stereo-video/>, composed

of 17 min of low-bit rate encoded stereoscopic video); each of these two corpora combines indoor/outdoor, unstable and arbitrary lighting conditions, still and high motion scenes, etc. (cf. Appendix);

- A new 3D video disparity map is considered, and its watermarking potentiality is objectively benchmarked against state-of-the-art solutions in terms of transparency, robustness, and computational cost;
- The transparency evaluation is carried out on both subjective and objective basis; the former considers the ITU-R BT 500-12 [25] and BT 1438 [26] recommendations, while the latter is based on two pixel difference-based measures (PSNR and image fidelity (IF)) and on three correlation-based measures (NCC, structural content (SC), and structural similarity (SSIM));
- The robustness is objectively expressed by means of the watermarked detection BER against several classes of attacks, such as linear and nonlinear filtering, compression, and geometric transformations;
- The computational cost is estimated for each and every processing step involved in the watermarking chain.

4 Disparity maps for stereoscopic video watermarking

In its widest acceptance, a disparity map provides information about the coordinates at which similar blocks are located in two images (the so-called reference and target images). Computing a disparity map requires to design a rule specifying how the reference block is searched for in a given area of the target image and to define a similarity metric establishing whether or not a reference block matches a target block. By exploiting the spatio/temporal correlation between successive frames in 2D video, several disparity maps have already been advanced and proved their efficiency in various research fields, like compression, indexing, or segmentation. They generally assumed that the differences between the target and reference frames are homogeneous on the two directions. The exhaustive search algorithm [27] is widely used for block motion estimation in video coding in order to determine effective similarity while providing minimal error estimation. However, running the full search window, this algorithm requires a massive computation. Hence, several fast algorithms were developed to reduce the computation time. Zeng and Liou [7] advanced the NTSS algorithm for fast block-matching estimation and showed that the algorithm provides smaller motion compensation errors comparing to the state of the art. The stereoscopic video compression was approached by the MPEG community via shape-adaptive block matching based on exhaustive (full) search, further referred to as FS-MPEG [8]: the block matching starts from 9×9 pixels and is progressively reduced (e.g., 5×5 or 3×3 in [9]) until the similarity measure reaches its local minimal value. During

Table 1 A comparative view on the state of the art of stereoscopic image/video-watermarking methods

View-based	Embedding technique	Corpus	Transparency marked vs. original views	Robustness recovered vs. original watermark	Data payload (per frame)	Estimated computational cost	Advantages	Drawbacks
View-based	SS in the DCT domain [16]	25 frames of 256×256 pixels	Average PSNR 34.89 dB	Average PSNR 19.04 dB	Visual logo 32×32 -pixel binary image	DCT computation $O(n \log(n))$ SS insertion $O(m)$	Objective transparency larger than 30 dB Low complexity	No subjective evaluation of transparency The attacks are not precise
	SS in the DWT domain [17]	3 frames of 512×512 pixels	Average PSNR 38.81 dB	Average PSNR 16.68 dB	Visual logo 64×64 -pixel binary image	DWT computation $O(n \log(n))$ SS insertion $O(m)$	Objective transparency larger than in [16] Low complexity	No subjective evaluation of transparency The attacks are not precise
	SS in the SVDDWT domain [18]	5 images of 512×512 pixels	Average PSNR 44.05 dB	Average NCC (normalized cross correlation) Average filtering, 0.91 Rotation, 0.91; resizing, 0.90; Gaussian noise, 0.89	Disparity map 128×128 -pixel image	DWT computation $O(n \log(n))$ SVD computation $O(\min(m^2, n, mn^2))$ SS insertion $O(m)$ Genetic algorithm	Objective transparency larger than in [16, 17]	Experiments on 3 frames No subjective evaluation of transparency Experiments on 5 frames
	SS in the FrFT domain [19]	2 frames of 256×256 pixels 1 frame of 512×512 pixels	Average PSNR 48.5 dB	Average NCC JPEG compression, 0.98; Gaussian noise, 0.46; average filtering, 0.71; rotation, 0.63; resizing, 0.69; cropping, 0.57	Disparity map 64×64 -pixel image 128×128 -pixel image	FrFT computation $O(n \log(n))$ SS insertion $O(m)$	Objective transparency larger than [16–18] Robustness to a larger class of attacks than in [18]	Very high complexity (genetic algorithm) No subjective evaluation of transparency Experiments on 3 frames
	QIM in the DWT domain [20]	NA	NA	Bit error rate (BER) Gaussian, 0.07; median, 0.11; row/column removal, 0.38; 0.25° rotation, 0.30; 0.5° rotation, 0.39	Random generated bits 2,000 bits	DWT computation $O(n \log(n))$ QIM insertion $O(m)$	Robustness evaluated against a larger class of attacks than in [16–18]	No evaluation of transparency Low robustness No information about the processed corpus
	SS in the combined DWT-DCT domain [21]	1 frame of 640×480 pixels	Average PSNR 52.07 dB	Watermark recovering rate (WRR) JPEG compression, 0.94; salt and pepper noise, 0.90; median filtering, 0.94; cropping, 0.81	Visual logo 64×64 -pixel binary image	DWT-DCT computation $O(n \log(n))$ SS insertion $O(m)$	The best transparency among the investigated view-based watermarking methods	No subjective evaluation of transparency Experiments on 1 frames
Disparity-based	SS in the disparity map [22]	3 video sequences of 22 frames of 512×512 pixels 1 video sequence of 192 frames of 720×576 pixels	Average degradation index 0.06 dB	WRR Recoding, 0.86; recoding and noise addition, 0.81	NA	Disparity map $O(n^2)$ SS insertion $O(m)$	The insertion in the disparity map	No subjective evaluation of transparency Robust only against recoding

this matching procedure, the vertical disparity component is assumed to be zero.

All these studies achieve the block matching by measures based on the differences between the values in the two blocks (e.g., mean square error (MSE) or sum of absolute differences (SAD)). Such approaches can be suboptimal for 3D video.³ First, the disparities between the left and the right views are predominant on the horizontal direction (where the very depth information is conveyed) while the vertical disparities are mainly connected to the video-shooting errors (the so-called rig errors). Secondly, basic metrics like MSE or SAD are unrelated to the human visual system and would a priori lack in achieving visually good block matching. These two challenges are addressed by the 3DV-NTSS method [28, 29].

4.1 The 3DV-NTSS disparity map

The general idea of the 3DV-NTSS algorithm is to follow the NTSS procedure while taking into account the spatial constraints brought by the stereoscopic context and the human visual system peculiarities. As the right and left cameras are located on the same horizontal plane (given by the rig), the horizontal disparities are to be preponderantly considered in depth computation, while vertical disparities are mainly linked to the rig alignments errors. Hence, the 3DV-NTSS algorithm assigns discriminative weights for the horizontal and vertical disparities which are subsequently used to adapt the vertical and horizontal sizes of the search area according to the content on which the disparity is estimated. The block matching is achieved by the NCC.

The 3DV-NTSS algorithm [28, 29] is illustrated in the flow chart in Fig. 3.

4.2 Quantitative results

The experiments consider three algorithms (NTSS, FS-MPEG, and 3DV-NTSS), each of which being individually run with two different matching similarity measures (SAD and NCC), on the two corpora presented in the Appendix.

Table 2 presents the quality of the reconstructed images. In this respect, two pixel difference-based measures (PSNR and IF) and three correlation-based measures (NCC, SC, and SSIM) have been considered [30, 31]. Note that the values reported in Table 2 are individually computed on each view in the reconstructed video sequence and subsequently averaged over all the frames in that corpus. In Table 2, the values corresponding to the state-of-the-art NTSS and FS-MPEG with SAD similarity metric are presented in bold while the 3DV-NTSS computed with the NCC similarity metric are presented in italics. The following conclusions can be drawn:

³ More details for 3D video disparity map computation can be found in [28, 29].

3DV-NTSS vs. NTSS

The 3DV-NTSS leads on the 3DLive corpus to relative gains of 6 % in PSNR, 7 % in IF, 12 % in NCC, 0 % in SC, and 6 % in SSIM. When considering the MPEG corpus, 3DV-NTSS and NTSS provide quite similar results in terms of PSNR and SC (average relative gains lower than 1 %) while 3DV-NTSS outperforms NTSS with gains of 4 % in IF, 2 % in NCC, and 2 % in SSIM. Note that these gains are computed as relative gains between the advanced 3DV-NTSS method based on the NCC block-matching criterion and the state-of-the-art reference given by the NTSS method based on the SAD block-matching criterion. For example, in the 3DLive and PSNR case, the relative gain of 6 % was computed as $(34.10 - 32.17) / 32.17 = 0.0593 \cong 6\%$.

3DV-NTSS vs. FS-MPEG

When considering the 3DLive corpus, the joint use of search area adaptivity and visual quality-based similarity metric in 3DV-NTSS leads to relative gains of 7 % in PSNR, 8 % in IF, 12 % in NCC, 1 % in SC, and 9 % in SSIM. On the MPEG corpus, these relative gains become 4 % in PSNR, 9 % in IF, 8 % in NCC, 2 % in SC, and 6 % in SSIM. Note that these gains are computed as explained above.

The computational cost of the considered algorithms was expressed as the number of search points per 16×16 block (see Table 3). The same three algorithms (NTSS, FS-MPEG, and 3DV-NTSS) and two matching similarity measures (SAD and NCC) have been considered. The values in Table 3 show that, when compared to the NTSS and FS-MPEG algorithms, the 3DV-NTSS features a computational cost decreased by a factor between 1.3 and 13, on both the 3DLive and MPEG corpora.

Tables 2 and 3 point to the good performances of the 3DV-NTSS disparity map (for more details in this respect, see [28, 29]). Consequently, it will be further considered as one of the potential watermarking insertion domains.

5 Watermarking benchmarking

A general view on the watermarking assessment procedure carried out in our study is presented in Fig. 4.

Four insertion methods are considered; they belong to the SS, SI (binary QIM and five-symbol QIM) and hybrid (IProtect) families.

Each of these four methods is successively applied on the left view of the video sequences as well as on three disparity maps, computed according to the NTSS, FS-MPEG, and 3DV-NTSS algorithms (all of them considering the NCC as the block-matching criterion). In our study, the mark insertion actually takes place in the 2D-DWT representation of each of these four insertion domains (see Fig. 5); however,

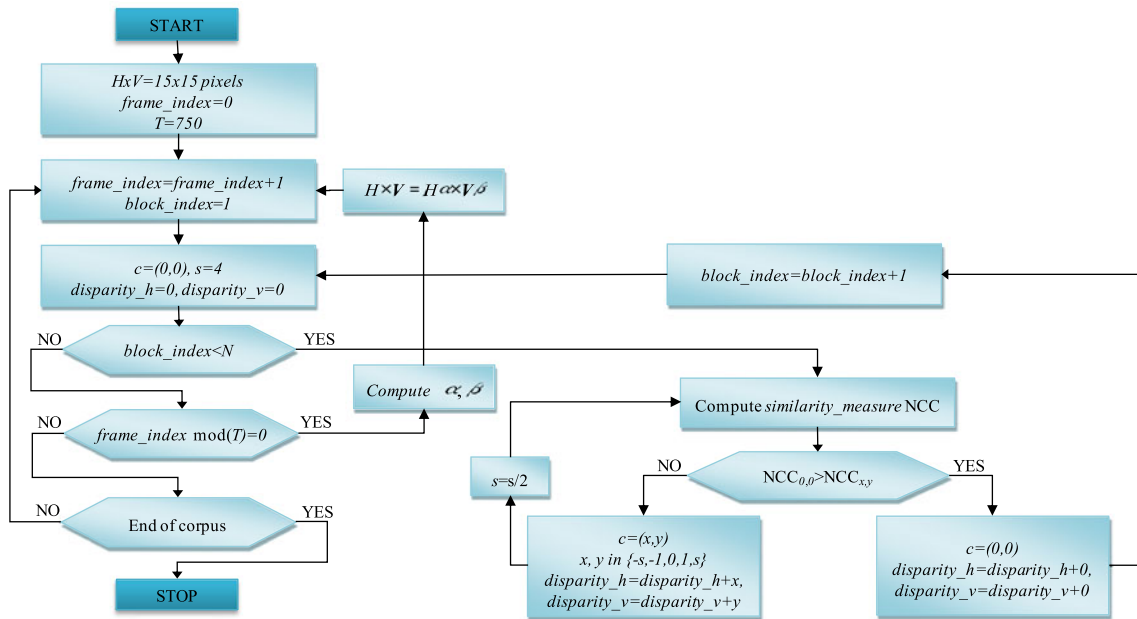


Fig. 3 The 3DV-NTSS algorithm flow chart

for simplicity, these domains will be further referred to as left view, NTSS, FS-MPEG, and 3DV-NTSS.

In the experiments, the *2D-DWT* is applied at an $N_r=6$ resolution level for $1,920 \times 1,080$ -pixel images, at an $N_r=4$ for 640×480 -pixel image, and at $N_r=3$ for images smaller than 320×192 pixels. The watermark is embedded in the hierarchy of the coefficients belonging to the HL_{N_r} and LH_{N_r} sub-bands.

The watermarking properties are evaluated in terms of transparency, robustness, and computational cost.

The transparency is assessed by both subjective and objective procedures. The former relies on the ITU-R BT 500-12 [25] and BT 1438 [26] recommendations and concerns the image quality, the depth perception, and the visual comfort. The latter is performed based on five objective image quality metrics, namely, PSNR, IF, NCC, SC, and SSIM. In all the transparency evaluation experiments, the method parameters were set so as to ensure a fixed robustness, expressed by a BER lower than 0.05 for JPEG compression attack.

The robustness is assessed by computing the BER in the watermark detection after five types of attacks, namely, the Gaussian filtering, sharpening, JPEG compression, and geometric (small rotations and StirMark random bending) attacks. These attacks are selected so as to represent the main classes of attacks mentioned by the DCI standards. In all the robustness tests, the method parameters were set so as to ensure a fixed transparency expressed by a PSNR of 30 dB.

The computational cost is not only expressed by the computational time needed to insert the mark but also by an analysis of the computation complexity.

All the experiments considered both the 3D Live and the MPEG stereoscopic video corpora (cf. Appendix).

The quantity of inserted information is kept unchanged in all the cases, namely, 1 bit per frame (i.e., 25 bits per second); note that this value is 200 times larger than the lower limit imposed by the DCI standards (35 bits per 5 min of video).

Table 2 Reconstructed image visual quality, expressed by PSNR, IF, NCC, SC, and SSIM. Each time (for each disparity map and each video corpus), two similarity measures have been considered for block matching: SAD and NCC

ALGORITHM		PSNR (dB)		IF		NCC		SC		SSIM	
		SAD	NCC	SAD	NCC	SAD	NCC	SAD	NCC	SAD	NCC
3D Live	NTSS	32.17	32.38	0.931	0.939	0.876	0.922	0.998	0.999	0.937	0.991
	FS-MPEG	31.96	32.24	0.918	0.939	0.873	0.916	1.011	1.008	0.912	0.968
	3DV-NTSS	33.58	34.08	0.953	0.995	0.908	0.979	0.989	0.999	0.988	0.997
MPEG	NTSS	32.56	33.32	0.684	0.743	0.889	0.887	1.025	1.004	0.917	0.938
	FS-MPEG	31.30	32.60	0.654	0.743	0.844	0.867	1.046	0.995	0.880	0.902
	3DV-NTSS	31.51	32.50	0.702	0.711	0.876	0.909	1.012	1.021	0.907	0.931

Table 3 Computational cost expressed in number of search points per 16x16 block. Each time (for each disparity map and each video sequences), two similarity measures have been considered for block matching: SAD and NCC

	3DLive			MPEG		
	NTSS	FS-MPEG	3DV-NTSS	NTSS	FS-MPEG	3DV-NTSS
SAD	25.07	222.36	16.85	22.48	216.09	16.80
NCC	21.77	222.36	16.85	21.60	216.09	16.80

5.1 Transparency evaluation

5.1.1 Subjective protocol

During the subjective testing procedure, the ITU-R BT 500-12 [25] and BT 1438 [26] recommendations were followed.

The evaluation has been conducted at two locations: in professional testing conditions at Cesson Sévigné 3D theater and in laboratory conditions at the ARTEMIS Department. In the latter case, a 47-in. LG LCD, full HD 3D monitor (1,920×1,080 pixels), and a 400-cd/m² maximum brightness is used in the experiments.

The test was conducted on a total number of 25 nonexpert viewers (hence, larger than the ITU-R lower limit set at 15), with marginal knowledge on the image quality. The age distribution ranges from 20 to 50, with an average of 25. All the subjects are screened for visual acuity using Snellen chart and color vision using the Ishihara test [26].

The experiments involved two subjects per session. The subjects are seated in line with the center of the monitor, at a distance of 2 m which is about the height of the screen multiplied by a factor of 3. Each observer evaluates 34 randomly chosen video excerpts of 40 s each. These excerpts represent the two corpora and all the possibilities investigated in the experiments: original video content and watermarked video content, obtained through four methods applied on four insertion domains.

A double stimulus continuous quality scale method has been adapted. The visual comfort, image quality, and depth perception are scored on a quality scale with five levels going from 1 to 5 (bad, poor, fair, good, and excellent; see Table 4). For the result analysis, the mean opinion score (MOS) is computed for each test condition as the average of the individual score.

The experimental results concerning the image quality, the depth perception, and the visual comfort are synoptically presented in Figs. 6, 7, and 8, respectively, where the MOS values are displayed alongside their 95 % confidence intervals [32].

These three figures are organized in the same way. The abscissa is decrementally divided into three levels in order to represent all the investigated cases. First, the two corpora (3DLive and MPEG) are figured out on the left and right sides, respectively. Secondly, for each corpus, the four watermarking methods (SS, 2-QIM, 5-QIM, and IProtect) are presented from left to right. Finally, for each method, each of the four insertion domains (left view, NTSS, FS-MPEG, and 3DV-NTSS) is depicted from left to right. The ordinate gives the MOS values (represented in squares) and the 95 % confidence limits in its estimation (represented in vertical lines centered on the related MOS values). The 95 % confidence limits obtained when evaluating the original content are also presented (in horizontal continuous red lines).

The values reported in Figs. 6, 7, and 8 allow us to formulate the following general conclusions:

- The 3DLive watermarked content results in better visual experience than the MPEG watermarked content, with an average MOS difference of 0.16 (this average value is computed for a given corpus, over all the insertion methods and all the insertion domains and for the image quality, depth perception and visual comfort scores). Such a behavior can be explained by the difference in the quality of the original contents which may influence the viewers in assessing the quality of the watermarked content.
- When comparing among the four classes of watermarking methods, it can be noticed that IProtect offers the best visual quality, with an average MOS larger by 0.18 in the

Fig. 4 The watermarking assessment procedure

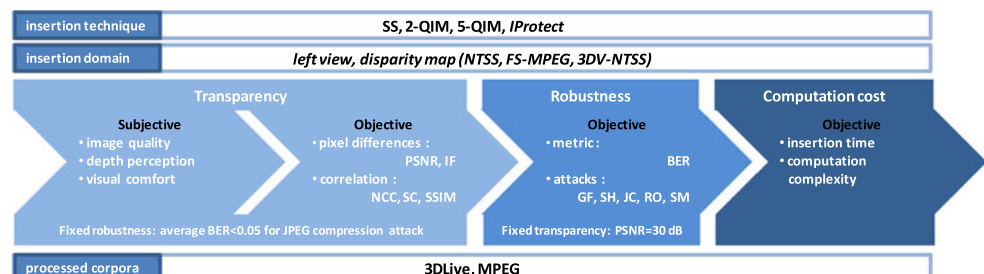
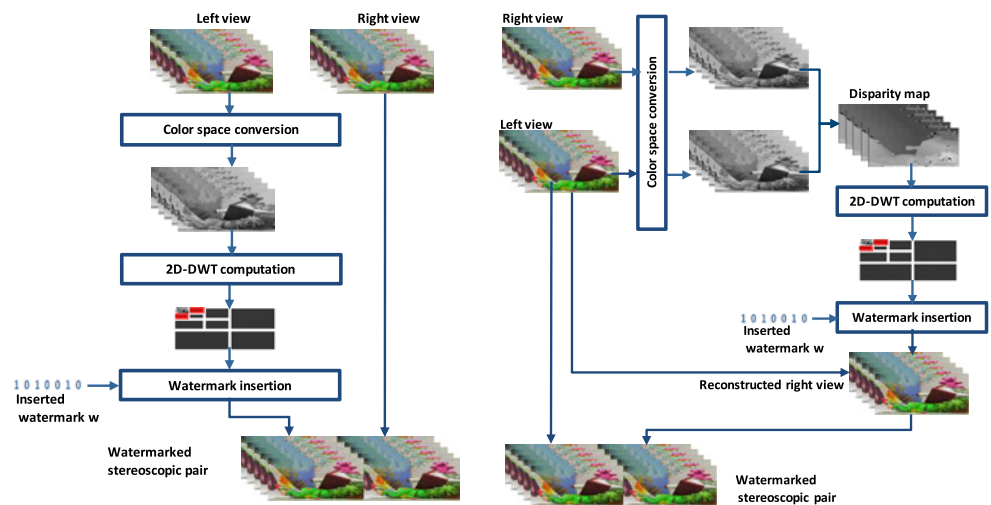


Fig. 5 Watermark insertion synopsis



(a) View-based scheme.

(b) Disparity-based scheme.

3DLive case and by 0.11 in the MPEG case with respect to the SS (this average value is computed for a given corpus and the corresponding insertion method, over all the insertion domains and for the image quality, depth perception, and visual comfort scores). SS is the second best choice. This result is unexpected, as for the 2D video content, the SS methods were reported to have the best visual quality [1, 2].

- When comparing among the four insertion domains, it can be noticed that 3DV-NTSS offers the best visual quality, with an average MOS larger by 0.06 in the 3DLive case and by 0.12 in the MPEG case, with respect to the NTSS (these average values of 0.06 and 0.12 are computed for a given corpus and the corresponding insertion domain, over all the insertion methods and for the image quality, depth perception, and visual comfort scores). NTSS is the second better domain. This result enforces the usefulness of the 3DV-NTSS disparity map for watermarking applications.
- The IProtect applied in the 3DV-NTSS domain is the only solution for achieving visually imperceptibly watermarking insertion. Actually, the IProtect/3DV-NTSS is the only combination ensuring for both corpora (3DLive and MPEG) and for the three evaluation criteria (image quality, depth perception, and visual comfort) confidence

limits inside the confidence limits corresponding to the original content. There is only one exception (the 3DLive corpus and the visual comfort) for which the lower limit of the IProtect/3DV-NTSS confidence interval is outside the confidence limits of the original content; even in this case, the IProtect/3DV-NTSS provides the better results.

Note that the conclusions above are meant to be general. However, several types of methods/insertion domains may be alternatively considered in order to solve particular applications defined by a particular type of content/targeted quality criterion/accepted quality limits. For instance, the protection of some low-quality MPEG content can be achieved under the depth perception constraints by three types of solutions: SS/3DV-NTSS, IProtect/FS-MPEG, and IProtect/3DV-NTSS (cf. Fig. 7).

5.1.2 Objective assessment

The visual quality of the watermarked content is objectively evaluated by five objective measures, namely, the PSNR, IF, NCC, SC, and SSIM. For each watermarking method and insertion domain, each of these five measures is first computed at the view level and then averaged at the corpus level. Figures 9, 10, 11, 12, and 13 represent the corresponding average values. These figures are organized in the same way as Figs. 6, 7, and 8.

The 95 % confidence limits are also computed for each experiment (each corpus, watermarking method, and insertion domain). As the corresponding error is always lower than 0.45 dB in the PSNR case and 0.001 in the IF, NCC, SC, and SSIM cases, the related confidence limits cannot be presented in Figs. 9, 10, 11, 12, and 13, being lower than the printing resolution for the average values. There is only one exception: for the 3DLive corpus, the IProtect method and the view-based domain, the 95 % error in SC estimation is 0.015 and is presented in Fig. 12.

Table 4 Computation complexity of the watermarking modules involved in IProtect/3DV-NTSS

IProtect/3DV-NTSS step	Data size	Complexity
2D-DWT	$n=h \times w$	$O(n \log(n))$
3DV-NTSS disparity map computation	$K=n/(Bs \times Bs)$	$K \times O(\log(p+1))$
IProtect insertion	$N \times m$	$N \times O(m)$
Image reconstruction	$n=h \times w$	$O(n)$
2D-IDWT computation	$n=h \times w$	$O(n \log(n))$

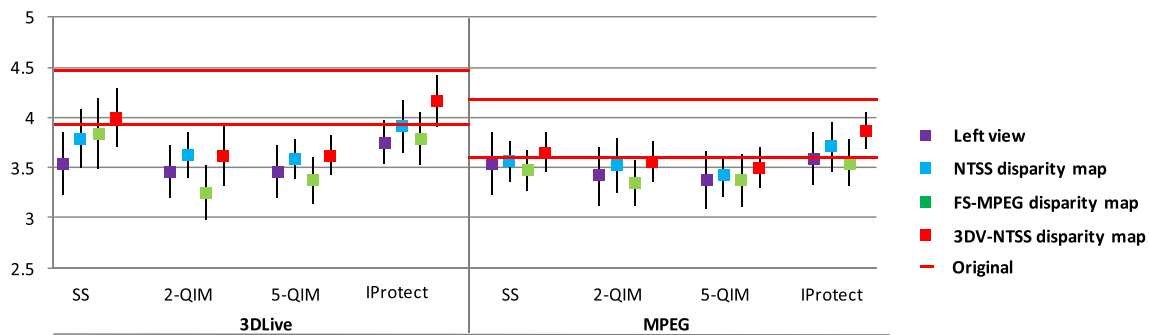


Fig. 6 Subjective evaluation for image quality: MOS values (in *squares*) and the related 95 % confidence limits (in *vertical lines* centered on the MOS) for watermarked content. The original content subjective evaluation is represented by its 95 % confidence limits (in *horizontal red lines*)

The values reported in Figs. 9, 10, 11, 12, and 13 allow us to formulate the following general conclusions:

- For a given measure and insertion domain, very few differences can be noticed between the 3DLive and MPEG corpora. This result validates the fairness of the benchmarking conditions (i.e., the parameters of the investigated watermarking methods were set so as to ensure a transparency independent with respect to the original data and dependent only on the method/insertion domain).
- The PSNR average values are always larger than 30 dB (with a singular exception, namely, the 5-QIM method applied to the left views of the 3DLive corpus). Consequently, all the considered watermarking methods/insertion domains can ensure basic transparency properties. However, very good transparency (larger than 35 dB) can be afforded only by SS (for all insertion domains and for the two corpora) and by IPProtect (3DV-NTSS in the 3DLive case and all the four insertion domains in the MPEG case). According to the PSNR values, SS would be the best watermarking method, followed by IPProtect.
- The IF, NCC, SC, and SSIM values also support the idea that basic transparency (i.e., values between 0.95 and 1.05) can be virtually ensured by all the considered watermarking methods (with some constraints in the choice of the insertion domain). There is one exception, represented by the SS method, which is refuted by the IF

measures estimated on the MPEG corpus. Here again, SS and IPProtect identified themselves as the best solutions.

- For each watermarking method and for each corpus, all the five objective quality metrics select the 3DV-NTSS disparity map as the optimal insertion domain, with a singular exception (the NCC values computed for the 5-QIM insertion method applied to the 3DLive corpus).

5.2 Robustness evaluation

During the experiments, three types of attacks are considered: filtering (Gaussian and sharpening), JPEG compression, and geometric (both small rotations and StirMark random bending). The Gaussian filtering and the sharpening consider a 3×3 convolution kernel. The JPEG compression is applied at a $Q=60$ quality factor. The rotations are randomly generated with angles lower than $\pm 0.5^\circ$. The StirMark random bending is applied at its default parameters [33, 34].

The robustness is evaluated by the BER in the watermark detection after the attacks. Each attack is individually applied at the view/disparity map level, and then, the corresponding number of errors is averaged at the corpus level in order to compute the BER. These average values and their related 95 % confidence limits are reported in Figs. 14, 15, 16, 17, and 18, which are organized in the same way as Figs. 6, 7, and 8.

When inspecting the results reported in Figs. 14, 15, 16, 17, and 18, it can be noticed that IPProtect/3DV-NTSS is the only combination ensuring a robustness expressed by a BER lower

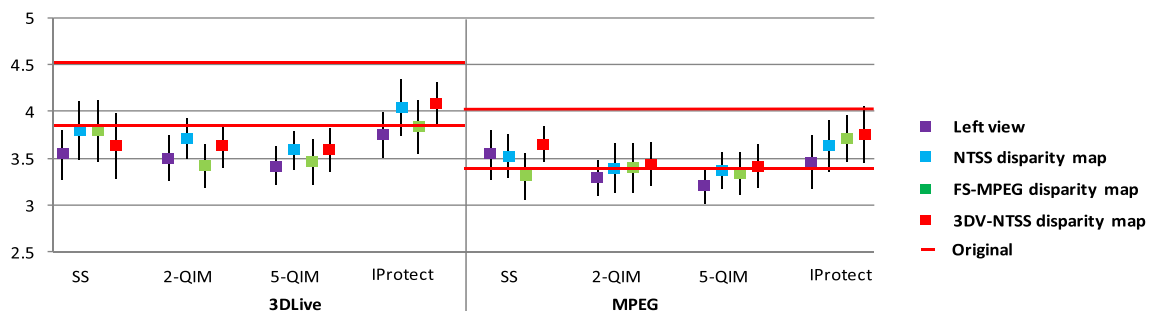


Fig. 7 Subjective evaluation for depth perception: MOS values (in *squares*) and the related 95 % confidence limits (in *vertical lines* centered on the MOS) for watermarked content. The original content subjective evaluation is represented by its 95 % confidence limits (in *horizontal red lines*)

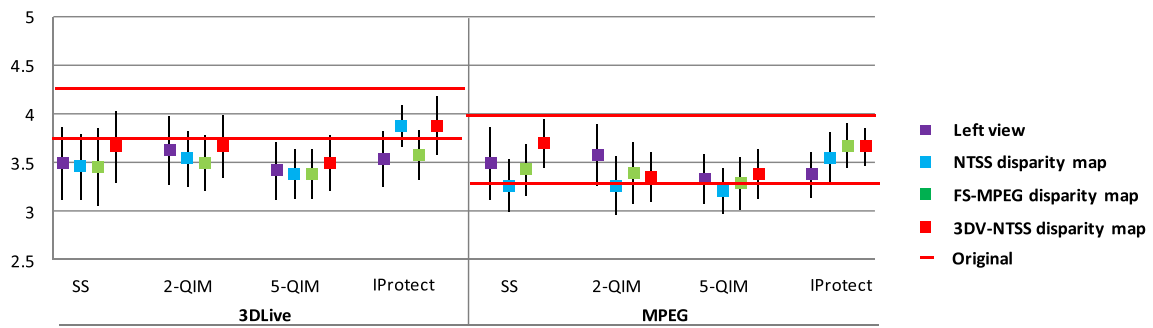


Fig. 8 Subjective evaluation for visual comfort: MOS values (in *squares*) and the related 95 % confidence limits (in *vertical lines* centered on the MOS) for watermarked content. The original content subjective evaluation is represented by its 95 % confidence limits (in *horizontal red lines*)

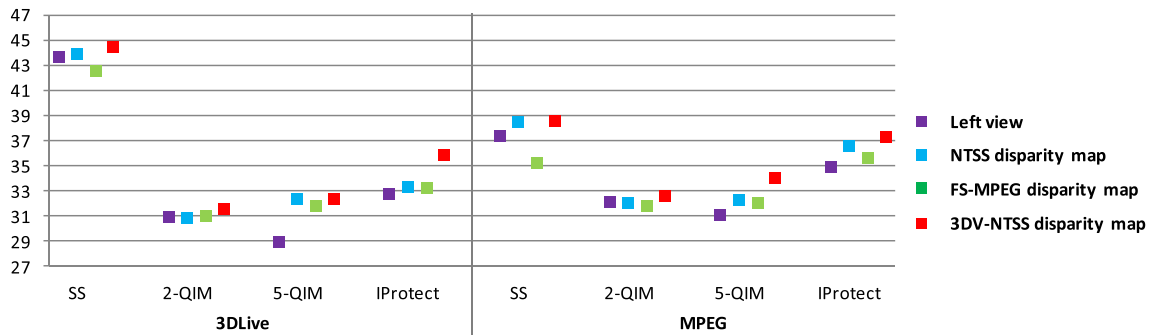


Fig. 9 Objective evaluation of the watermarked content visual quality: PSNR average values

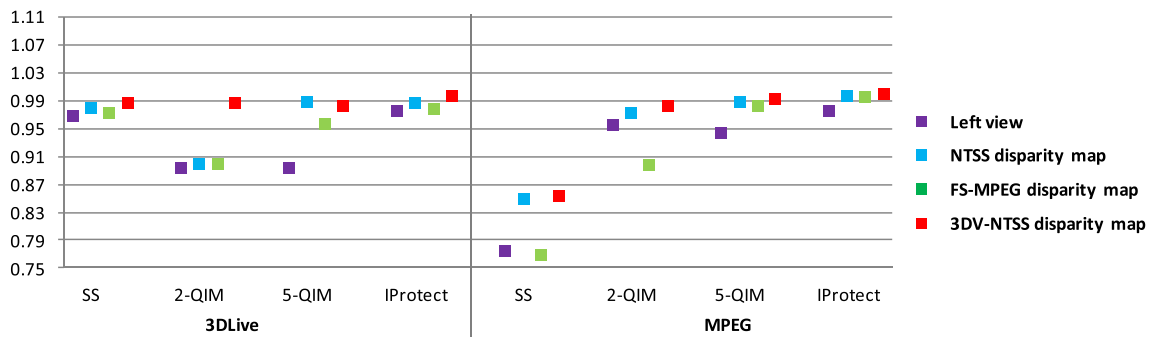


Fig. 10 Objective evaluation of the watermarked content visual quality: IF average values

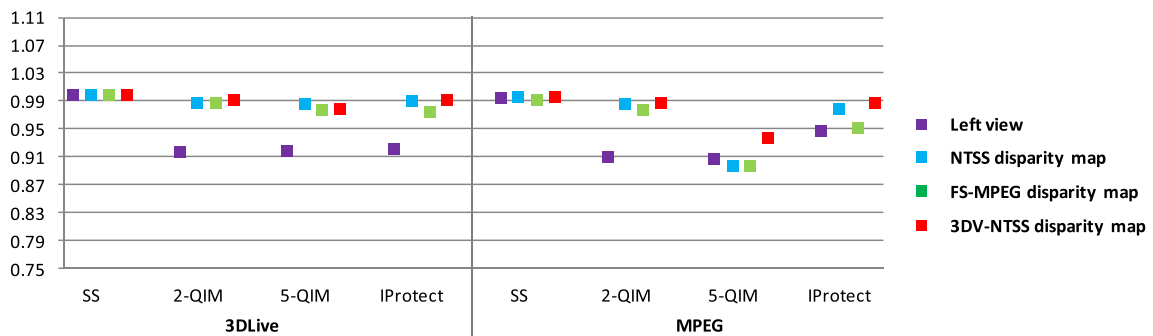


Fig. 11 Objective evaluation of the watermarked content visual quality: NCC average values

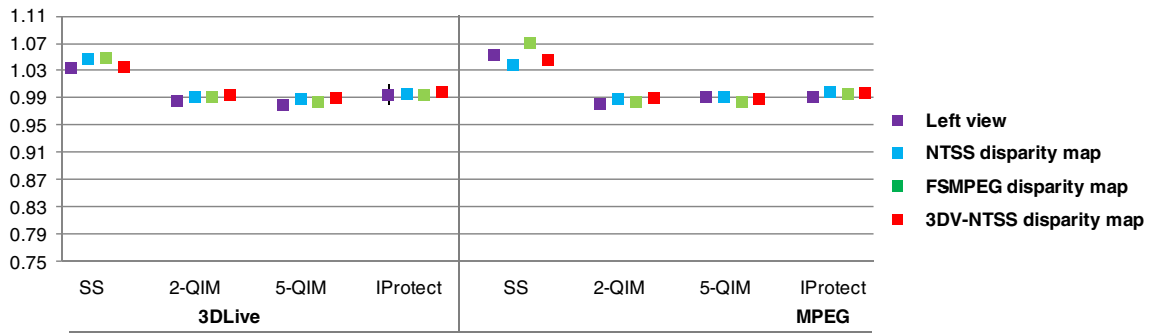


Fig. 12 Objective evaluation of the watermarked content visual quality: SC average values

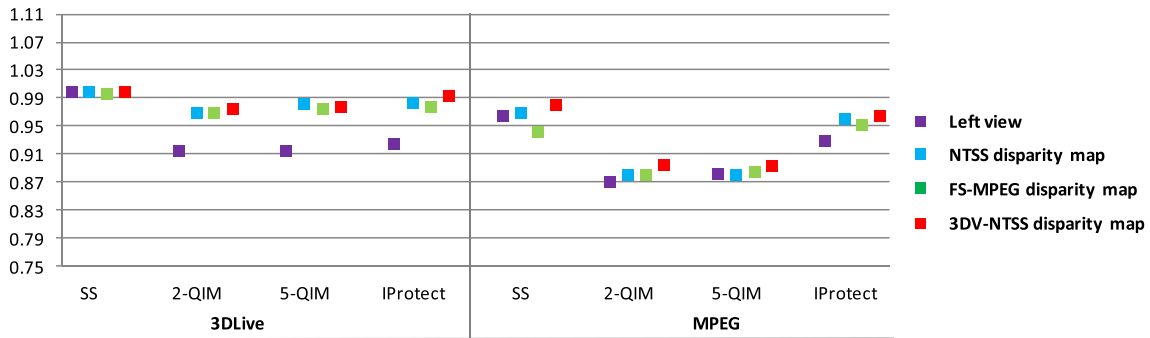


Fig. 13 Objective evaluation of the watermarked content visual quality: SSIM average values

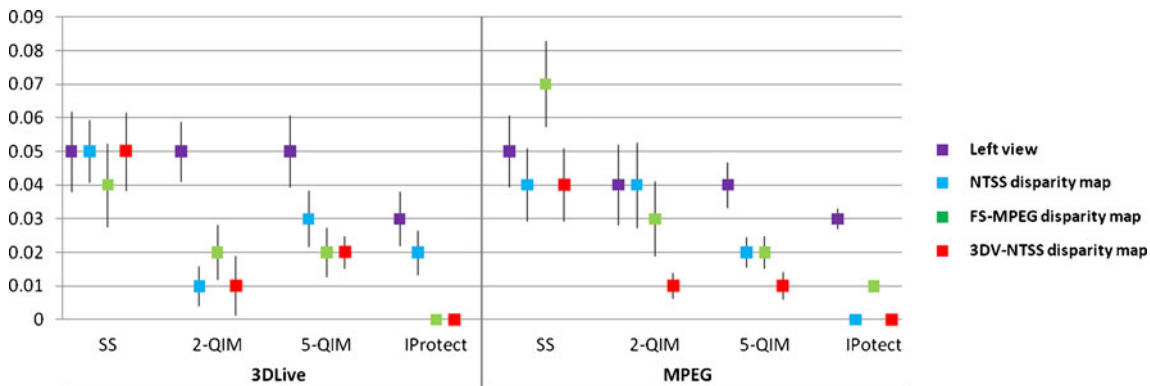


Fig. 14 Watermark robustness against Gaussian filtering: BER average value and the related 95 % confidence limits

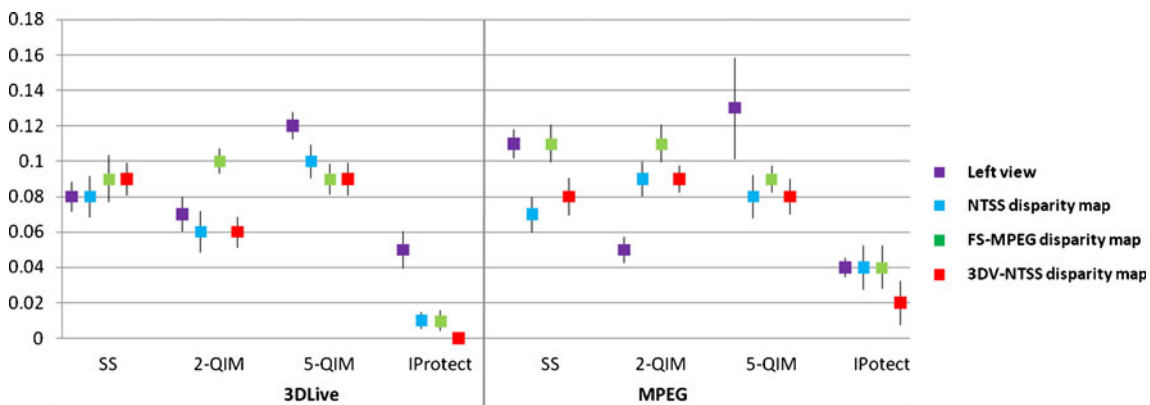


Fig. 15 Watermark robustness against sharpening: BER average value and the related 95 % confidence limits

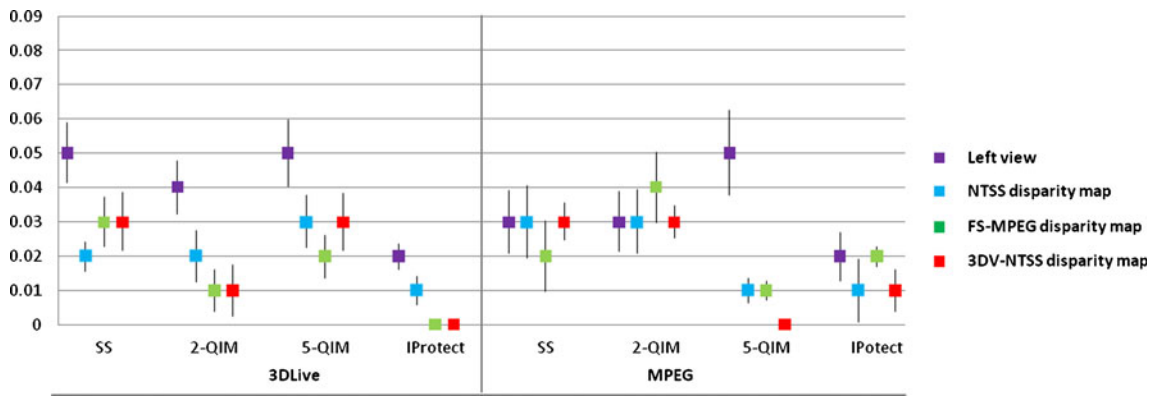


Fig. 16 Watermark robustness against JPEG compression: BER average value and the related 95 % confidence limits

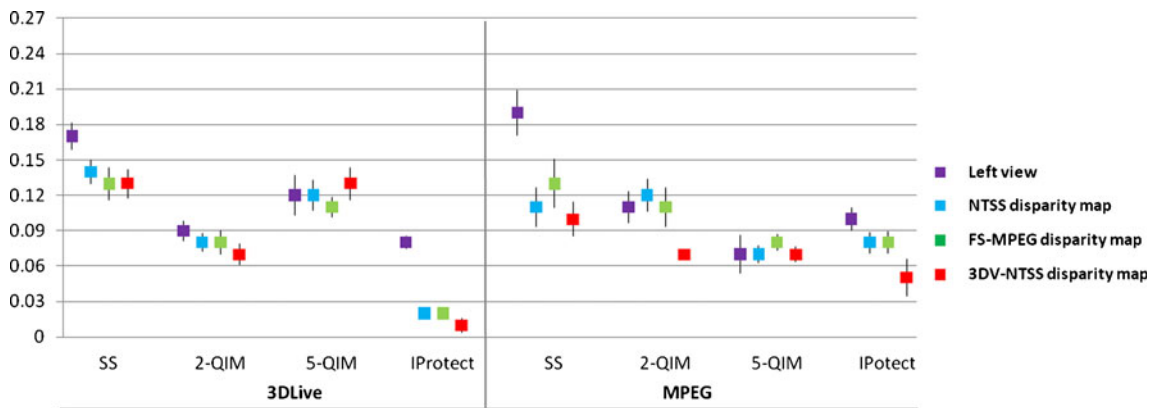


Fig. 17 Watermark robustness against small rotations: BER average value and the related 95 % confidence limits

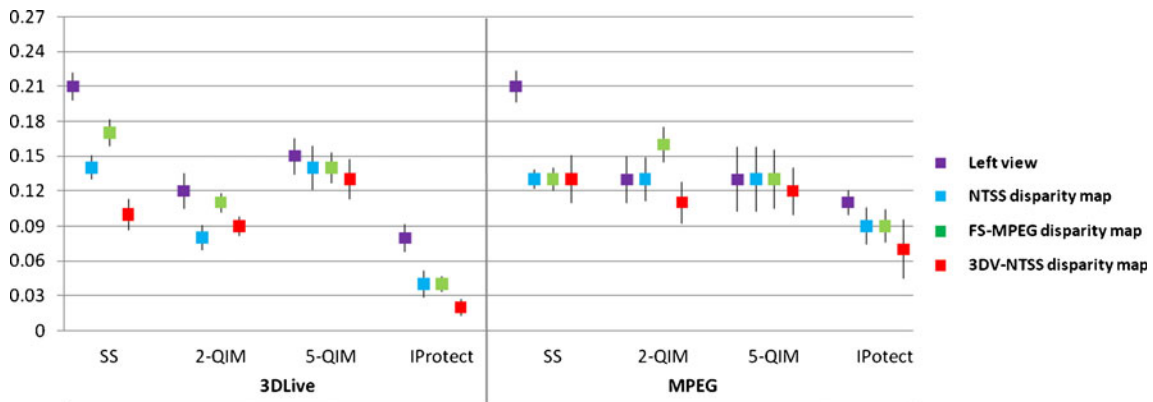


Fig. 18 Watermark robustness against geometric attacks (StirMark random bending): BER average value and the related 95 % confidence limits

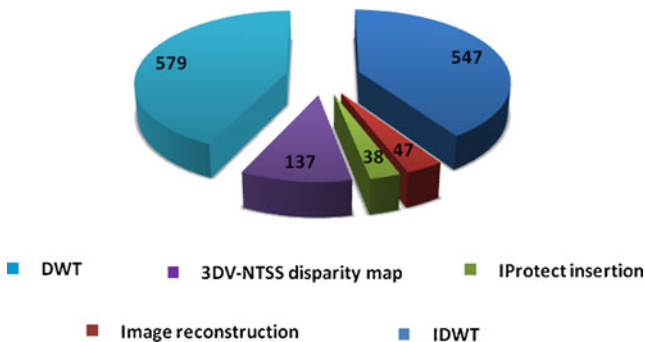


Fig. 19 Watermarking chain processing time (in ms)

than 0.05 after filtering and JPEG compression and lower than 0.1 after the geometric attacks, irrespective to the corpus.

The 2-QIM method applied on the 3DV-NTSS domain features the same good values of the BER against Gaussian filtering, compression, and rotations but fails in meeting the robustness requirements against sharpening (on both 3DLive and MPEG corpora) and against StirMark random bending (only in the case of the MPEG corpus).

Also, note that the SS method does not succeed in meeting the robustness requirements, irrespective to the insertion domain and/or processed corpus.

5.3 Computational cost

The results presented in the previous subsections hint to the IProtect watermarking method performed in the 3DV-NTSS disparity map as the most effective solution for stereoscopic video protection, when considering transparency and robustness constraints, for a fixed data payload. However, for several real-life applications (e.g., live HD 3D TV content protection), the computational cost should also be investigated.

In this respect, we evaluate the processing time and the computational complexity corresponding to the main five main steps in the IProtect/3DV-NTSS chain: 3DV-NTSS disparity map computation, DWT, IProtect mark insertion, inverse DWT, and image reconstruction based on the 3DV-NTSS.

The processing time evaluation is carried out at the frame level and then averaged at the corpus level. The numerical values presented in Fig. 19 are expressed in milliseconds; they are obtained on a PC Core2 CPU at 2.13 GHz and with 2 GB of RAM.

It can be noticed that the 3DV-NTSS, the watermark insertion, and the image reconstruction are 4, 15, and 12 times faster than a DWT computation, respectively. Consequently, for real-life solutions implementing the DWT/IDWT in real time [35], the IProtect/3DV-NTSS watermarking can be also carried out in real time. This is not the case of our software implementation, which should be considered in the present study only as an investigation tool.

The processing time intrinsically depends on the particular hardware and software experimental setup. A contrario, the computational complexity of the underlying algorithm can give an a priori estimation of the computational cost. The computational complexity values corresponding to the five main steps of the IProtect/3D-NTSS are presented in Table 4, where the following notations are made: h and w are the height and width of the frame for which the wavelet is computed, m is the watermark size, N is the iteration number required by the IProtect algorithm, and p is the pixel searching distance in block of size $B_s \times B_s = 16 \times 16$ pixels.

The values reported in Table 4 strengthen the conclusions drawn from the processing time analysis: the most complex steps in the IProtect/3DV-NTSS chain correspond to the direct/inverse wavelet transform computation.

6 Conclusion

This paper reports on a comparative study on the possibility of using 2D inherited watermarking methods for stereoscopic video protection. The comparison is carried out on the watermarking method (belonging to the SS, IE, and hybrid SS-IE classes) and on its underlying insertion domain (left view, NTSS, FS-MPEG, and 3DV-NTSS).

The experimental results brought to light the fact that IProtect (a hybrid spread spectrum–side information method) performed in a new disparity map domain (3DV-NTSS) would be generic enough to as to serve a large variety of applications:

- It is the only analyzed insertion method ensuring the imperceptibility of the watermarking according to the subjective tests (TU-R BT 500-12 and BT 1438 recommendations) performed so as to evaluate the image quality, the depth perception, and the visual comfort;
- This subjective transparency evaluation is reinforced by the objective assessment, exhibiting PSNR >35 dB and IF, NCC, SC, and SSIM limits within the (0.95, 1.05) interval;
- it is the only investigated method ensuring robustness expressed by BER lower than 0.05 after linear/nonlinear filtering and JPEG compression and lower than 0.1 after the geometric attacks; note that although the experiments considered only these four classes of attacks, the IProtect method is intrinsically (see Fig. 2) robust against any attack which can be modeled by an additive Gaussian noise (from noise addition to frame-synchronous re-encoding); the time-variant frame jitter induced by some particular video re-encoding operations can be counterattacked by a content-based resynchronization module inserted prior to the detection, as designed in [36] for fingerprinting purposes.
- These practical watermarking performances can be achieved with a nonprohibitive computational cost, compatible to the real-time applications (in the sense discussed in Section 5.3).

Note that such results are made possible by the 3DV-NTSS method which is designed so as to take advantage of both the inner relations between the left and right views and of a human visual system-related measure in the block-matching algorithm.

The generality of the results is ensured by the size and composition of the two corpora (a total of 2 h, 28 min, and 53 s of heterogeneous stereoscopic video content) and by the statistical error control (95 % confidence limits) in all the results.

Future work will be devoted to the reconsideration of the IProtect for protecting DIBR stereoscopic data watermarking [37, 38].

Acknowledgments The experimental validation was performed with in the 3DLive French project.

Appendix

All the experiments reported in the present paper are carried out on two corpora, further referred to as 3DLive and MPEG. Each of these two corpora combines indoor/outdoor, unstable and arbitrary lighting, and still and high motion scenes, as illustrated in Figs. 20 and 21, respectively.

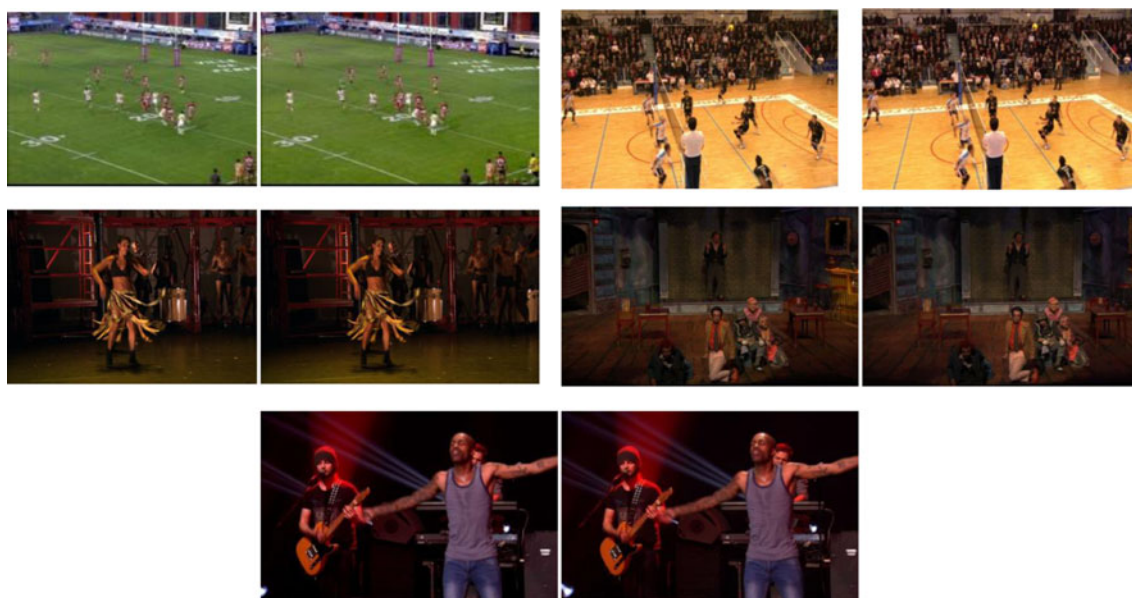


Fig. 20 Left and right views sampled from the 3DLive corpus: rugby, volley, dancing, theater, and rock band sequences



Fig. 21 Left and right views sampled from the MPEG corpus: rollerbalade, office, city tours, and cartoon sequences

Organized under the framework of the 3DLive French project, the 3DLive corpus (<http://3dlive-project.com>) sums up 2 h, 11 min, and 24 s of stereoscopic video sequences (197,000 stereoscopic pairs encoded at 25 frames per second). It regroups content sampled from a rugby match, a volleyball match, dancing/theater performances, and a private gig of a rock band. These sequences are full HD-encoded ($1,920 \times 1,080$ pixels).

The MPEG corpus (<http://sp.cs.tut.fi/mobile3dtv/stereo-video/>) is composed of 41 sequences and sums up to 17 min and 29 s (29,908 stereoscopic pairs, as several frame rates have been considered). Various resolutions are represented, from 320×192 to 640×480 pixels. The content corresponds to street events, like roller and biking races, indoor (office) scenes, city tours, cartoons, etc. Note that for the sake of uniformity, the frames in Figs. 20 and 21 are presented at the same size, although their actual sizes are very different, as explained above.

References

1. Cox I, Klian J, Leighton FT, Shamoon T (1997) Secure spread spectrum watermarking for multimedia. *IEEE Trans Image Process* 6(12)
2. Cox I, Miller M, Bloom J (2002) *Digital watermarking*. Morgan Kaufmann, San Mateo
3. Shannon CE (1958) Channel with side information at the transmitter. *IBM Journal of Research and Development* 2(4):289–293
4. Eggers JJ, Bäuml R, Tzschoppe R, Girod B (2003) Scalar Costa scheme for information embedding. *IEEE Trans Signal Process* 51(4)
5. Costa MHM (1983) Writing on dirty paper. *IEEE Trans Inf Theory* IT-29(3):439–441
6. Mitrea M, Prêteux F, Nunez J (2007) “Procédé de Tatouage d’une Séquence Vidéo.” French patent no. 05 54132 (December 2005); European extension under the number 1804213 (July 2007)
7. Li R, Zeng B, Liou ML (1994) A new three-step search algorithm for block motion estimation. *IEEE Trans Circ Syst Video Technol* 4(4):438–442

8. ISO/IEC JTC1/SC29/WG11 (2009) Depth estimation reference software (DERS) with image segmentation and block matching, M16092 (February 2009)
9. Stankiewicz O, Wegner K (2008) Depth map estimation software, ISO/IEC JTC1/SC29/WG11 MPEG/M15175, Antalya (January 2008)
10. Hasnaoui M, Belhaj M, Mitrea M, Prêteux F (2011) mQIM principles for MPEG-4 AVC watermarking, SPIE Photonics West 2011 (January 2011)
11. Belhaj M, Mitrea M, Duta S, and Preteux F (2010) MPEG-4 AVC robust video watermarking based on QIM and perceptual masking. In: International conference on communication, Bucharest (June 2010)
12. Mitrea M, Prêteux F (2007) Tatouage robuste des contenus multimédias. In: Chaouchi H, Laurent-Maknavicius M (eds) La sécurité dans les réseaux sans fil et mobiles. Lavoisier, Paris
13. Chammem A, Mitrea M, Preteux F (2011) DWT-based stereoscopic image watermarking. In: Proc. SPIE, vol 786326 (January 2011)
14. Lin S, Costello DJ Jr (1983) Error control coding: fundamentals and applications. Prentice-Hall, Englewood Cliffs
15. Mitrea M, Prêteux F, Petrescu M, Vlad A (2005) The StirMark watermarking attack in the DWT Domain. In: Proc. of the 12th international workshop on systems, signals and image processing, vol 2, Chalkida, Greece
16. Dong-Choon H, Kyung-Hoon B, Eun-Soo K (2003) Real-time stereo image watermarking using discrete cosine transform and adaptive disparity maps. Multimedia Systems Appl VI, Proc SPIE 5241:233
17. Dong-Choon H, Kyung-Hoon B, Eun-Soo K (2003) Stereo image watermarking scheme based-on discrete wavelet transform and adaptive disparity estimation. Mathematics of data/image coding, compression, and encryption, with applications. Conference no. 6, San Diego
18. Kumar S, Raman B, Thakur M (2009) Real coded genetic algorithm based stereo image watermarking. IJSDA Int J Secure Digit Inf Age 1(1)
19. Bhatnager G, kumar S, Raman B, Sukavanam N (2009) Stereo coding via digital watermarking. J Electron Imaging 18(3)
20. Campisi P (2008) Object-oriented stereo image digital watermarking. J Electron Imaging 17(4):043024
21. Yu M, Wang A, Luo T, Jiang G, Li F, Fu S (2011) New block relationship based stereo image watermarking algorithm. In: The sixth international conference on systems and networks communications, ICSNC 2011, pp 171–174 (October 2011)
22. Zhang Z, Zhu Z, Xi L (2007) Novel scheme for watermarking stereo video. Int J Nonlinear Sci 3(1):74–80
23. <http://3dlive-project.com>
24. <http://sp.cs.tut.fi/mobile3dtv/stereo-video/>
25. ITU-R (2002) Methodology for the subjective assessment of the quality of television pictures. Tech Rep. BT.500-11, ITU-R
26. ITU-R (2000) Subjective assessment of stereoscopic television pictures. Tech Rep BT.1438
27. Kuang K, Hosur PI (1999) Core experimental results of fast motion estimation based on new test conditions (Q4a), ISO/IEC JTC1/SC29/WG11/M4934, Vancouver, B.C. (July 1999)
28. Chammem A, Mitrea M, Prêteux F (2012) Adaptive disparity map computation for stereoscopic video watermarking. In: The IEEE international conference on electronics, circuits, and systems (ICECS), Seville, Spain (December 2012)
29. Chammem A, Mitrea M, Prêteux F (2012) High-definition three-dimensional television disparity map computation. J Electron Imaging 21(4):043024–043024
30. Eskicioglu MA, Fisher PS (1995) Image quality measures and their performance. IEEE Trans Commun 43(12)
31. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4)
32. Walpole RE, Myers RH, Myers SL, Ye K (2002) Probability & statistics for engineers and scientists. Pearson Educational International, Upper Saddle River
33. Petitcolas F, Ross JA, Markus GK (1998) Attacks on copyright marking systems. In: Proceedings on information hiding, second international workshop, IH'98, LNCS 1525, Springer, Portland, pp 219–239 (April 1998). ISBN 3-540-65386-4
34. Petitcolas F (2000) Watermarking schemes evaluation. IEEE Signal Process 17(5):58–64
35. Galiano V, Lopez O, Malumbres MP, Migallon H (2011) Improving the discrete wavelet transform computation from multicore to GPU-based algorithms. In: Proceedings of the 11th international conference on computational and mathematical methods in science and engineering, CMMSE 2011 (June 2011)
36. Garboan A, Mitrea M, Prêteux F (2013) Cinematography sequences tracking by means of fingerprinting techniques. Ann Telecommun 68(3–4):187–199
37. Franco-Conteras J, Baudry S, Doerr G (2011) Virtual view invariant domain for 3D video blind watermarking. In: IEEE international conference on image processing, pp 2817–2820
38. Koz A, Çiğla C, Aydın AA (2010) Watermarking of free view video. IEEE Trans Image Process 19:1785–1797