RR-03-080

Switching Between Orthogonal Watermarks for Enhanced Security Against Collusion in Video

G. Doërr and J.-L. Dugelay

July 4, 2003

Abstract

Recent efforts for watermarking digital video basically extend the results previously obtained for still images. Thus, most of the proposed algorithms rely on a frame-by-frame approach. Such a direct adaptation leads to non-secure algorithms because of collusion attacks. Depending of the targeted application, this may be a critical concern. A new watermarking system, extending current schemes based on spread spectrum, is consequently introduced in this paper. It basically embeds alternative orthogonal watermarks in successive video frames, which succeeds in obtaining superior performances in terms of security. Finally, a discussion is conducted to check that this new architecture cannot be easily broken with a straightforward attack.

Chapter 1

Introduction

Digital watermarking has now been investigated for ten years. This technology basically embeds a robust and invisible watermark in digital data. The watermark encodes the message associated with the targeted application: rights related to the data for copyright protection, client signature for traitor tracing, data signature for authentication... The embedded signal is inherently tied to the content and survives D/A conversion. Digital watermarking is indeed often regarded as a second line of defense once digital data has been left in clear after decryption. Sooner or later, encrypted digital data has to be decrypted to be presented to a human observer/listener. As a result, encryption does not protect any more the data and digital watermarking has been introduced here to fill this *analog gap*.

Digital watermarking, the art of hiding information in a robust and invisible manner, has consequently been investigated as a complementary technology. There exists a complex trade-off between the three parameters *data payload*, *fidelity* and *robustness* in digital watermarking. The data payload is the amount of information, i.e. the number of bits encoded by the hidden watermark. The fidelity is another property of the watermark: the distortion, which the watermark embedding process is bound to introduce, should remain imperceptible to a human observer. Finally, the robustness of a watermarking scheme can be seen as the ability of the detector to extract the underlying watermark from some altered watermarked data. Those parameters are conflicting and a compromise has to be found, which is often tied to the targeted application. For further insight regarding digital watermarking, the reader is redirected towards existing books [16, 5]. If digital watermarking has been mostly devoted to still images at the beginning, watermarking other types of multimedia data is currently investigated.

One of those *new objects* of interest is digital video. There are indeed many applications in the context of video [8] where inserting a digital watermark might be of interest. Cinema studios are reluctant to disseminate their high valued videos, which might be perfectly copied and rapidly distributed at large scale, and are requesting copyright protection services. The recently popular peer-to-peer networks are great tools to efficiently find and exchange digital data. However, a malicious user can also use it to distribute copyrighted data that he/she does not own. Traitor tracing should be introduced. Advertisers want to get what they have paid for i.e. to have their advertisements broadcasted during the air time they have booked. The broadcasts need consequently to be monitored in an automatic and reliable way. In few words, many applications can benefit from digital watermarking. However it would be utopian to think that a single watermarking system will fit them all. Depending on the targeted application, the specifications will be slightly different, particularly in terms of security i.e. resistance of the watermark against hostile intelligence (collusion).

To date, video watermarking has mostly inherited from the results obtained for still images. If some algorithms exploit the specificities of a compression standard [14, 20] or embed a watermark in a three dimensional transform [6, 27], watermarking digital video content is regarded most of the time as watermarking a sequence of still images. The drawback of such a straightforward adaptation is that it does not consider the very specific nature of video content and this results in weak algorithms in terms of security. This may be critical depending on the targeted application. In Chapter 2, two reference spread spectrum video watermarking schemes are described. Two collusion attacks are then introduced in Chapter 3 which succeed in trapping the previously introduced watermarking schemes. As a result, a new architecture is proposed in Chapter 4 which is proven to be robust against both attacks. In Chapter 5, the solutions left to the attacker are discussed to show that the system cannot be easily broken. Finally, conclusions are drawn and tracks for future work given in Chapter 6.

Chapter 2

Video Watermarking

One of the pioneer techniques for video watermarking has been described by Hartung and Girod [12]. It basically relies on the Spread Spectrum (SS) theory [25]. The payload to be hidden is first duplicated and then frequency spread thanks to a modulation with a pseudo-random noise. The resulting watermark signal is then scaled by an embedding strength and added to the video signal. This approach is still used as an underlying layer in recent video watermarking schemes [15, 23]. Depending on how the modulation is performed, two different systems can be obtained, which will be described in the next two subsections.

2.1 SS System

In the original algorithm [12] that Hartung and Girod proposed, video is considered as a one-dimensional signal and the modulation is global. In other terms, a different watermark is inserted in each video frame [23] as depicted in Figure 2.1. This approach will be referred as *SS system* in the remaining of the article and is further described below.

On the embedder side, a pseudo-random watermark is inserted in each video frame according to the following additive embedding rule:

$$\check{F}_t = F_t + \alpha W_t(K), \quad W_t(K) \sim \mathcal{N}(0,1) \tag{2.1}$$

where F_t is the t^{th} video frame, $\check{F}_t$ its watermarked version, α the embedding strength and K a secret key. The inserted watermark $W_t(K)$ has a normal distribution with zero mean and unit variance and should be different at every instant t. A simple way to obtain this property is to use $K + t$ as a seed for the pseudo-random generator. Perceptual shaping can be introduced to improve the invisibility of the watermark by making for example the embedding strength α dependent of the local content of the frame [29]. In practice, a global embedding strength has been used and its value has been set equal to 3, so that the resulting distortion is around 38 dB in terms of Peak Signal to Noise Ratio (PSNR).

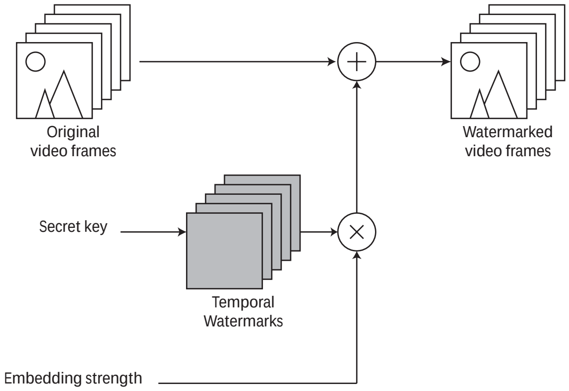

Figure 2.1: Description of SS system embedder.

On the detector side, the presence/absence of a watermark is checked thanks to a correlation score $\rho(.)$ computed as follows:

$$\begin{aligned}\rho\big(\{\check{F}_t\}\big) &= \frac{1}{T}\sum_{t=1}^{T}\check{F}_t \cdot W_t \\ &= \frac{1}{T}\sum_{t=1}^{T}F_t \cdot W_t + \frac{\alpha}{T}\sum_{t=1}^{T}W_t \cdot W_t \\ &= \alpha + \frac{1}{T}\sum_{t=1}^{T}F_t \cdot W_t \\ &\approx \alpha \qquad (2.2)\end{aligned}$$

where T is the number of considered video frames and $\cdot$ denotes the linear correlation operation. Other detection metrics can be used, like the normalized correlation or the correlation coefficient for example. Here linear correlation has been chosen to facilitate the derivations. Moreover, prefiltering can be performed before computing the correlation score so that the detection statistics are enhanced. This correlation score should be equal to α if a watermark is present in the video, while it should be almost equal to zero if no watermark has been inserted. As a result, the computed score is compared to a threshold τ_{detect} in order to assert the presence or the absence of the watermark. The value given to this detection threshold will determine the false positive and false negative probabilities and it need to be chosen carefully depending on the targeted application. In practice, the value $\alpha/2$ has been chosen which results in equal false positive and false negative probability.

2.2 SS-1 System

A major drawback of the previous system is that it relies on temporal synchronization on the detector side. As a result, a simple frame drop or insertion succeeds in confusing the detector. Consequently, one may rather insert the same watermark as depicted in Figure 2.2, encoding multiple bits, in each video frame. The frames are subsequently accumulated in a buffer before detection. This is basically what Philips has done for its video watermarking algorithm JAWS [15]. This approach will be referred as *SS-1 system* in the remaining of the article and is further described below.

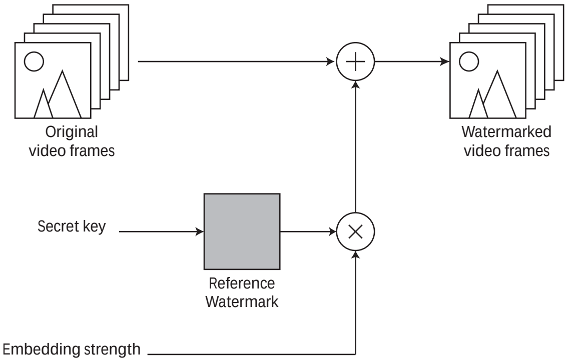

Figure 2.2: Description of SS-1 system embedder.

On the embedder side, the same pseudo-random watermark is inserted in each video frame with the additive embedding rule given in Equation 2.1. This can be written:

$$\check{F}_t = F_t + \alpha W_0(K), \quad W_0(K) \sim \mathcal{N}(0,1) \tag{2.3}$$

where $W_0(K)$ is a reference watermark which has a normal distribution with zero mean and unit variance and which has been pseudo-randomly generated with the secret key K used as a seed. Once again, a global embedding strength has been used and its value has been set equal to 3 to obtain a distortion around 38 dB in terms of PSNR.

On the detector side, the same correlation score $\rho(.)$ than in Equation 2.2 is computed in order to assert the presence or the absence of the watermark. However, since the same watermark is inserted in each video frame, the linearity of the operator $\cdot$ can be exploited as follows:

$$\rho\big(\{\check{F}_t\}\big) = \frac{1}{T}\sum_{t=1}^{T} \check{F}_t \cdot W_0 = \left(\frac{1}{T}\sum_{t=1}^{T} \check{F}_t\right) \cdot W_0 \tag{2.4}$$

This equation means that averaging several correlations between different video frames and the same watermark is equivalent to computing a single correlation between the average of the video frames and this watermark. This reduces the number of computations required for the detection. Here again, the correlation score should be equal to α if a watermark is present in the video, while it should be almost equal to zero if no watermark has been inserted. As a result, the computed score is compared to a threshold τ_{detect}, which is set equal to $\alpha/2$ in practice, in order to assert the presence or the absence of the reference watermark W_0.

Chapter 3

Collusion Attacks

The robustness of a watermarking scheme is often evaluated via the survival of the embedded watermark after attacks. This has motivated the development of efficient benchmarking tools [2] to compare different watermarking systems. However, video watermarking algorithms have not yet been evaluated in a hostile environment. In other terms, the verification process simply checks if the watermark survives attacks without any underlying malicious intelligence, e.g. noise addition, frame filtering, chrominance resampling (4:4:4, 4:2:2, 4:2:0), lossy compression, transcoding, changes across display formats (4/3, 16/9, 2.11/1), changes of spatio-temporal resolution (NTSC, PAL, SECAM), etc. Nevertheless, digital watermarks protecting high-valued video items are likely to be submitted to strong hostile attacks when they are distributed to a large audience. Basically, several watermarked contents can be colluded to produce unprotected content [26]. Collusion traditionally occurs when a clique of malicious customers gather together to produce unwatermarked content. That is *inter-videos* collusion i.e. several watermarked video are required to produce unprotected content. Additionally, successive frames of a watermarked video can be regarded as several watermarked images. Thus, a single malicious user can collude several watermarked frames to produce an unprotected video. That is *intra-video* collusion i.e. a watermarked video alone permits to stir out the watermark signal from the video stream. The next two subsections describe two of such intra-video collusion attacks, which succeed in removing a watermark inserted by one or the other previous systems.

3.1 Temporal Frame Averaging

Digital watermarks are generally localized mostly in high frequencies since the Human Visual System (HVS) is less sensible to noise addition. As a result, one of the earliest proposed attack to remove hidden watermarks is to apply a low-pass filter to the protected data [19]. Spatial filtering has been investigated extensively and most watermarking algorithms for still images are robust against it today. In the context of video, since neighbor video frames are highly similar, temporal low-pass filtering can be used to obtain an estimate of the original video frames i.e. without the underlying watermark. This can be written:

$$\dot{F}_t = \mathcal{L}_w(E_t), \quad E_t = \{F_u, -w/2 \leq t-u < w/2\} \tag{3.1}$$

where w is the size of the temporal window, $\mathcal{L}_w$ is the used temporal low-pass filter and $\dot{F}_t$ is the resulting t^{th} attacked video frame. In practice, a simple temporal averaging filter has been used with a window size w equal to 3. The attack will be considered as a success if the correlation score computed with the attacked video frames falls below the detection threshold τ_{detect}.

Proposition 1 *Averaging successive video frames succeeds (resp. fails) in removing a digital watermark embedded in a video with the SS system (resp. SS-1 system).*

Proof When a watermarked video $\{\check{F}_t\}$ is temporally averaged, the resulting attacked video frames are given by:

$$\begin{aligned} \dot{F}_t &= \frac{1}{w} \sum_{u \in [-\frac{w}{2}, \frac{w}{2}[} \check{F}_{t+u} \\ &= \frac{1}{w} \sum_{u \in [-\frac{w}{2}, \frac{w}{2}[} F_{t+u} + \frac{\alpha}{w} \sum_{u \in [-\frac{w}{2}, \frac{w}{2}[} W_{t+u} \end{aligned} \tag{3.2}$$

As a result, when the correlation score is computed on the detector side, the following result is obtained:

$$\begin{aligned} \rho\big(\{\dot{F}_t\}\big) &= \frac{1}{T} \sum_{t=1}^{T} \dot{F}_t \cdot W_t \\ &= \frac{1}{wT} \sum_{t=1}^{T} \Bigg(\sum_{u \in [-\frac{w}{2}, \frac{w}{2}[} F_{t+u} \cdot W_t \Bigg) \\ &\quad + \frac{\alpha}{wT} \sum_{t=1}^{T} \Bigg(\sum_{u \in [-\frac{w}{2}, \frac{w}{2}[} W_{t+u} \cdot W_t \Bigg) \\ &\approx \frac{\alpha}{wT} \sum_{t=1}^{T} \Bigg(\sum_{u \in [-\frac{w}{2}, \frac{w}{2}[} W_{t+u} \cdot W_t \Bigg) \end{aligned} \tag{3.3}$$

Depending on the strategy enforced during embedding, the correlation score is reduced or not. With the SS-1 system, the same watermark W_0 is embedded redundantly i.e. $W_t = W_0$ for all t. Consequently, every correlation term $W_{t+u} \cdot W_t$ is equal to 1 and the correlation score is equal to α. In other terms, temporal frame averaging has no effect on a watermark inserted by the SS-1 system. Alternatively, with the SS system, watermarks inserted in neighbor frames are uncorrelated. As a result, only the term corresponding to the index $u = 0$ in the summation $\sum_{u \in [-\frac{w}{2}, \frac{w}{2}[} W_{t+u} \cdot W_t$ will not be null. This results in a final correlation score equal to α/w. It means that, when the watermark has been inserted with the SS system, temporal frame averaging with a window size w reduces the correlation score by a factor w. For w greater than 2, it implies that the attack has made the correlation score drop below the detection threshold τ_{detect} which is equal to $\alpha/2$. ▮

In summary, if the same watermark has been embedded in all the video frame, averaging several frames is completely useless since it has no impact on the detector. On the contrary, enforcing an *always insert a different watermark* strategy on the embedder side introduces a weakness against temporal frame averaging on the detector side. This result has to be contrasted with the content of the video scene. Indeed, averaging several successive frames may result in a poor quality video if fast moving objects are present in the scene or if there is a camera global motion. As a result, this attack is particularly relevant in static scenes.

3.2 Watermark Estimation Remodulation

When all the video frames carry the same watermark, the attacker can estimate the embedded watermark in each video frame and obtain a refined estimation of the watermark by combining (e.g. taking the average) those different estimations.

$$\tilde{W}_T = \frac{1}{T}\sum_{t=1}^{T}\mathcal{E}(\check{F}_t) \tag{3.4}$$

where $\mathcal{E}(.)$ is a watermark estimator. The ideal estimator consists in computing the difference between a watermarked video frame and the associated original one as follows:

$$\mathcal{E}_0(\check{F}) = \check{F} - F \tag{3.5}$$

where F is an original video frame and $\check{F}$ its watermarked version. However, in practice, an attacker has not access to the original video frames and the watermark estimation process should be done in a blind manner. Previous work [28] has been done to estimate a watermark inserted in an image. As previously mentioned, a digital watermark is generally localized in high frequencies and a quite good estimation can be obtained by computing the difference between a watermarked video frame and its low-pass filtered version:

$$\mathcal{E}_1(\check{F}) = \check{F} - \mathcal{L}(\check{F}) \tag{3.6}$$

where $\mathcal{L}(.)$ is a spatial low-pass filter. Unfortunately, in practice, some sample are badly estimated e.g. around the edges of the frame and in textured regions. As a result, an additional thresholding operation is performed to remove those non-pertinent samples:

$$\mathcal{E}_2(\check{F}) = \mathcal{T}_{\tau_{estim}}\Big(\check{F} - \mathcal{L}(\check{F})\Big) \tag{3.7}$$

where the thresholding function $\mathcal{T}_\tau(.)$ is defined as follows:

$$\mathcal{T}_\tau(x) = \begin{cases} x & \text{if } |x| < \tau \\ 0 & \text{otherwise} \end{cases} \tag{3.8}$$

In practice, a simple 5×5 spatial averaging filter is used for $\mathcal{L}(.)$ and the threshold τ_{estim} has been set to 14. Moreover, fifty successive video frames are used to estimate the potentially embedded watermark.

Once the embedded watermark has been estimated, the attacker can subtract it in each watermarked video frame $\check{F}_t$ with a strength β. This process can be written:

$$\dot{F}_t = \check{F}_t - \beta \tilde{W}_T \tag{3.9}$$

where $\dot{F}_t$ is the resulting attacked version of the t^{th} video frame. Again, perceptual shaping can be introduced during this step by making the remodulation strength β dependent of the local content of the frame. In practice, an attacker usually follows another constraint which is that the visual distortion introduced by the attack should be similar to the one introduced by the watermark embedding process. The Mean Square Error between an original video frame and its watermarked version is equal to $\alpha^{\,2}$, while the distortion between a watermarked video frame and its attacked version is equal to $\beta^2 \tilde{W}_T \cdot \tilde{W}_T$. Following the previously stated constraint, Equation 3.9 becomes:

$$\dot{F}_t = \check{F}_t - \alpha \frac{\tilde{W}_T}{\sqrt{\tilde{W}_T \cdot \tilde{W}_T}} = \check{F}_t - \alpha \tilde{W}_{T_N} \tag{3.10}$$

where $\tilde{W}_{T_N}$ is the estimated watermark after normalization. Once again, the attack will be considered as a success if the correlation score computed with the attacked video frames falls below the detection threshold τ_{detect}.

Proposition 2 *The watermark estimation remodulation attack succeeds (resp. fails) in removing a digital watermark embedded in a video with the SS-1 system (resp. SS system).*

Proof Let assume that the attacker has access to the ideal watermark estimator $\mathcal{E}_0(.)$ i.e. the attacker is in the best possible position. When a watermarked video $\{\check{F}_t\}$ is considered, the resulting estimation of the watermark is given by:

$$\tilde{W}_T = \frac{1}{T} \sum_{t=1}^{T} W_t \tag{3.11}$$

Subsequently, when this estimated watermark is remodulated, the following attacked video frames are obtained:

$$\dot{F}_t = F_t + \alpha \left(W_t - \frac{\tilde{W}_T}{\sqrt{\tilde{W}_T \cdot \tilde{W}_T}} \right) \tag{3.12}$$

When the detector checks the presence of the watermark in those attacked video frames, the correlation score is equal to:

$$\begin{aligned}
\rho\big(\{\dot{F}_t\}\big) &= \frac{1}{T}\sum_{t=1}^{T}\dot{F}_t\cdot W_t \\
&= \frac{1}{T}\sum_{t=1}^{T}F_t\cdot W_t+\frac{\alpha}{T}\sum_{t=1}^{T}W_t\cdot W_t \\
&\quad -\frac{\alpha}{T\sqrt{\tilde{W}_T\cdot\tilde{W}_T}}\sum_{t=1}^{T}\tilde{W}_T\cdot W_t \\
&\approx \alpha\left[1-\frac{1}{T^2\sqrt{\tilde{W}_T\cdot\tilde{W}_T}}\sum_{t=1}^{T}\sum_{u=1}^{T}W_u\cdot W_t\right. \qquad (3.13)
\end{aligned}$$

Now, depending on which embedding strategy has been enforced, the correlation score will be reduced or not. With the SS system, watermarks embedded in different video frames are uncorrelated and $W_u \cdot W_t = \delta_u^t$ where δ is the Kronecker delta. It means that only the term corresponding to the index $u = t$ contributes to the summation $\sum_{u=1}^{T} W_u \cdot W_t$. It also means that the norm $\sqrt{\tilde{W}_T \cdot \tilde{W}_T}$ of the estimated watermark is equal to $1/\sqrt{T}$. Combining those two results, it appears that the correlation score after the attack is equal to $\alpha(1 - 1/\sqrt{T})$ which is almost equal to α for large T. Since T is required to be quite large to obtain a good estimate, the attack is a failure. On the other hand, with the SS-1 system, the same watermark W_0 is embedded redundantly in all the video frames and the resulting estimated watermark $\tilde{W}_T$ is equal to W_0, whose norm is equal to 1. All the terms in the summation $\sum_{u=1}^{T} W_u \cdot W_t$ are equal to 1 and the final correlation score drops down to 0. In other terms, the watermark estimation remodulation attack completely removes a watermark embedded with SS-1 system. ■

To sum up, when a different watermark is embedded in every video frame, the watermark estimation remodulation attack is a failure. Indeed, the underlying idea behind this attack is that there is a single watermark redundantly embedded i.e. a single watermark to be estimated. Since there are many more in this context, the remodulation of the estimated watermark will have no significant impact on the correlation score. Alternatively, enforcing an *always insert the same watermark* strategy makes sense to the estimation process. There is a single watermark to be estimated and combining individual estimates should output a refined estimation of the redundantly inserted watermark. In fact, the more the video frames are different, the more each individual watermark estimate refines the final one. In other terms, this attack is more efficient in dynamic scenes. Once the secret hidden watermark has been estimated, the remodulation process completely removes the watermark from the different watermarked video frames and the attack is a success.

Chapter 4

Novel Approach

The danger of collusion is not always critical depending on the targeted application. In a broadcast monitoring context [15], it is not worthy for the broadcaster to remove a watermark from a video commercial for example. The advertiser will indeed detect that the commercial has not been aired and sue the television company in court. In a content authentication environment [24], if a watermark is removed, the carrier content will not be considered as reliable and immediately discarded. Nevertheless, there are many upcoming applications where collusion has to be addressed since it may open ways for forgery and later on result in a drastic loss of royalties. For example, in a Pay-Per-View or Video-On-Demand architecture, the service provider owns some high-valued video items, may be with exclusive rights. Some security devices need consequently to be inserted to prevent, for example, a customer who breaks his/her license agreement from capturing the video signal, burning it on a CD-ROM and distributing it widely. Digital watermarking may be a solution and a new system is described in Section 4.1. The proposed approach is then proven to be robust against intra-video collusion, first theoretically in Section 4.2 and then experimentally in Section 4.3.

4.1 SS-N System

One lesson learned from the previous watermarking systems is that embedding a single fixed watermark in all the frames of the video introduces a weakness against the watermark estimation remodulation attack. This has consequently motivated the design of a multi-watermarks scheme as depicted in Figure 4.1, which will be referred as *SS-N system* in the remaining of the article. On the embedder side, for each video frame, a watermark is randomly chosen from a finite set of N watermarks $\{W_i\}$. Each watermark has a normal distribution with zero mean and unit variance and has a probability p_i to be chosen at an instant t. Moreover, the watermarks are orthonormalized with the Gram-Schmidt algorithm [3] to prevent cross-talk on the detector side.

$$W_i \cdot W_j = \delta_i^j \quad \forall (i,j) \tag{4.1}$$

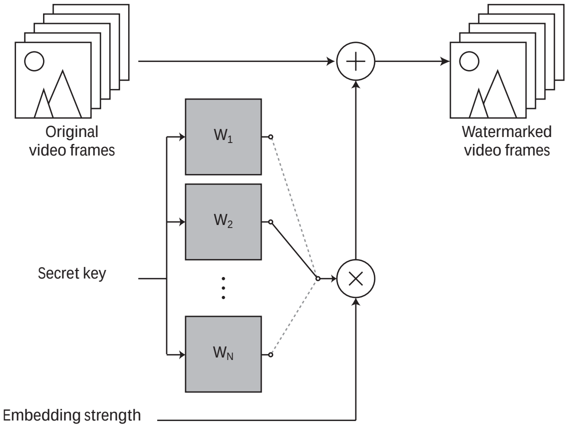

Figure 4.1: Description of SS-N system embedder.

The chosen watermark is then inserted in the video frame with the additive embedding rule given in Equation 2.1. This can be written:

$$\check{F}_t = F_t + \alpha W_{\phi(t)}, \quad P\big(\phi(t) = i\big) = p_i \tag{4.2}$$

where the p_i's can be regarded as the emission probabilities of the system. Here again, a global embedding strength equal to 3 has been used, which results in a distortion around 38 dB in terms of PSNR. It should be noted that both previous systems are a specific case of this novel architecture: $N = 1$ for SS-1 system and $N = \infty$ for SS system.

The second lesson learned with the previous systems is that, when different watermarks are embedded in successive video frames, temporal frame averaging with a window size w spreads the energy of a given watermark over the w neighboring frames. As a result, a new correlation score is computed on the detector side:

$$\rho\big(\{\check{F}_t\}\big) = \frac{1}{T}\sum_{t=1}^{T}\Big(\sum_{i=1}^{N}|\check{F}_t \cdot W_i|\Big) \tag{4.3}$$

In other terms, for each video frame, N linear correlation scores are computed in parallel and their absolute values are summed before being temporally averaged. The complexity of the detector has been increased by a factor N, which may prevent real-time detection when N grows large. Moreover, the absolute values prevents from using the linearity of the operator $\cdot$ to reduce the complexity as in Equation 2.4. Anyway, off-band detection remains interesting in many video applications e.g. traitor tracing. When detection is performed after embedding i.e. without any attack, the following correlation score is obtained:

$$\begin{aligned}\rho\big(\{\check{F}_t\}\big) &= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}|F_t\cdot W_i+\alpha W_{\phi(t)}\cdot W_i| \\ &\approx \frac{\alpha}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}|W_{\phi(t)}\cdot W_i| \\ &\approx \frac{\alpha}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\delta_{\phi(t)}^{\;i} \\ &\approx \alpha \end{aligned} \tag{4.4}$$

Several interfering terms $F_t\cdot W_i$ appear in the summation and the detection statistic can be improved by removing any correlation between the original video frames and the set of watermarks in a preprocessing step [4]. Once again, the correlation score should be equal to α if a watermark is present in the video, while it should be almost equal to zero if no watermark has been inserted. As a result, the computed score is compared to a threshold τ_{detect}, which is set equal to $\alpha/2$ in practice, in order to assert the presence or the absence of an hidden watermark.

4.2 Enhanced Security

The SS-N system has been proposed as a more secure alternative to the previously presented watermarking systems. It has indeed been shown that the SS system is weak against temporal frame averaging, while the system SS-1 is weak against the watermark estimation remodulation attack. Consequently, it will be checked in this subsection that this new system keeps its promise.

Proposition 3 *Temporal frame averaging does not remove a digital watermark embedded in a video with the SS-N system.*

Proof Let assume that a watermarked video $\{\check{F}_t\}$ is temporally averaged with a large window size w i.e. the attacker perform a really strong attack without any concern for the video quality. The resulting attacked video frames are then given by:

$$\begin{aligned}\dot{F}_t &= \frac{1}{w}\sum_{u\in[-\frac{w}{2},\frac{w}{2}]}\check{F}_{t+u} \\ &= \frac{1}{w}\sum_{u\in[-\frac{w}{2},\frac{w}{2}]}F_{t+u}+\frac{\alpha}{w}\sum_{u\in[-\frac{w}{2},\frac{w}{2}]}W_{\phi(t+u)} \\ &= \frac{1}{w}\sum_{u\in[-\frac{w}{2},\frac{w}{2}]}F_{t+u}+\alpha\sum_{i=1}^{N}p_iW_i \end{aligned} \tag{4.5}$$

Subsequently, when the detector compute the correlation score with those attacked video frames, the following result is obtained:

$$\begin{aligned}\rho\big(\{\dot{F}_t\}\big) &= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}|\dot{F}_t\cdot W_i| \\ &\approx \frac{\alpha}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\Big|\sum_{j=1}^{N}p_j\delta_j^i\Big| \\ &\approx \frac{\alpha}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}p_i \\ &\approx \alpha \qquad (4.6)\end{aligned}$$

Temporal frame averaging does not affect the correlation score and the attack is a success. $\blacksquare$

Here, the robustness against temporal frame averaging is achieved by checking the presence of *all the watermarks* of the set $\{W_i\}$ in each video frame. Since temporal frame averaging spreads the energy of a watermark embedded in a video frame over the neighbor frames contained in the temporal window, such a process permits to retrieve all the parts of each watermark. With the SS system, the detector only checks for the presence of the watermark that should be embedded in a video frame and consequently misses most of the watermark signal which is present in the neighbor frames.

Proposition 4 *The watermark estimation remodulation attack does not remove a digital watermark embedded in a video with the SS-N system.*

Proof Let assume again that the attacker has access to the ideal watermark estimator $\mathcal{E}_0(.)$ i.e. the attacker is in the best possible position. If a single watermark is estimated from a watermarked video $\{\check{F}_t\}$, the following result is obtained:

$$\tilde{W}_T = \frac{1}{T}\sum_{t=1}^{T}W_{\phi(t)} = \sum_{i=1}^{N}p_iW_i \qquad (4.7)$$

Later on, when this estimated watermark is remodulated, the produced attacked video frames look like:

$$\dot{F}_t = F_t + \alpha\left[\Big(1-\frac{p_t}{K}\Big)W_{\phi(t)} - \sum_{i\neq\phi(t)}\frac{p_i}{K}W_i\right] \qquad (4.8)$$

where $K = \sqrt{\tilde{W}_T\cdot\tilde{W}_T}$ is the norm of the estimated watermark. Now, if the detector checks the presence of the watermark in those attacked video frames, the computed correlation score is equal to:

$$
\begin{aligned}
\rho\big(\{\dot{F}_t\}\big) &= \frac{1}{T}\sum_{t=1}^{T}\Big(\sum_{i=1}^{N}|\dot{F}_t \cdot W_i|\Big) \\
&\approx \frac{\alpha}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\left|\Big(1-\frac{p_{\phi(t)}}{K}\Big)\delta_{\phi(t)}^{\;i} - \sum_{j\neq\phi(t)}\frac{p_j}{K}\delta_{\phi(t)}^{\;j}\right| \\
&\approx \frac{\alpha}{T}\sum_{t=1}^{T}\left[\Big(1-\frac{p_{\phi(t)}}{K}\Big) + \sum_{j\neq\phi(t)}\frac{p_j}{K}\right] \\
&\approx \alpha\sum_{i=1}^{N}p_i\left[\Big(1-\frac{p_i}{K}\Big) + \sum_{j\neq i}\frac{p_j}{K}\right] \qquad (4.9)
\end{aligned}
$$

Now let assume that all the p_i are equal to $1/N$. The norm K of the estimated watermark $\tilde{W}_T$ is then equal to $1/\sqrt{N}$. When this is introduced in the previous equation, the computed correlation score becomes:

$$
\rho\big(\{\dot{F}_t\}\big) = \alpha\left[1 + (N-2)\frac{\sqrt{N}}{N}\right] \qquad (4.10)
$$

In other terms, for N greater or equal to 2, the correlation score is greater or equal to τ_{detect} and the attack is a failure. ■

The robustness against this second attack is mainly due to the use of several watermarks, which confuses the watermark estimation process of the attacker. In each video frame, the attacker can only remove a small fraction $\sqrt{N}/N$ of the embedded watermark signal. On the other hand, he/she also removes a small part of all the other watermarks from the set $\{W_i\}$. Then, summing the absolute values of the linear correlations succeeds in compensating the loss of correlation with the originally embedded watermark. In fact, the reader can notice that the absolute values play a key role here. If the absolute values are removed from the correlation score formula, the algorithm is still immune to temporal frame averaging but the watermark estimation remodulation attack makes then the correlation score drop down to zero. Finally, one can also notice that for $N = 1$ (SS-1 system), the above formula predicts that the attack is a success.

4.3 Experimental Results

In order to verify the validity of the previous theoretical proofs, experiments have been conducted on several small video shots of 50 frames taken from five larger videos. The shots alternate between static and dynamic video content, as well as static and moving camera. The video shots are watermarked with the three previously presented watermarking schemes with an embedding strength equal to 3. Three different watermarks have been used for the SS-N system. The watermarked videos are then submitted to temporal frame averaging on one hand and to the watermark estimation remodulation

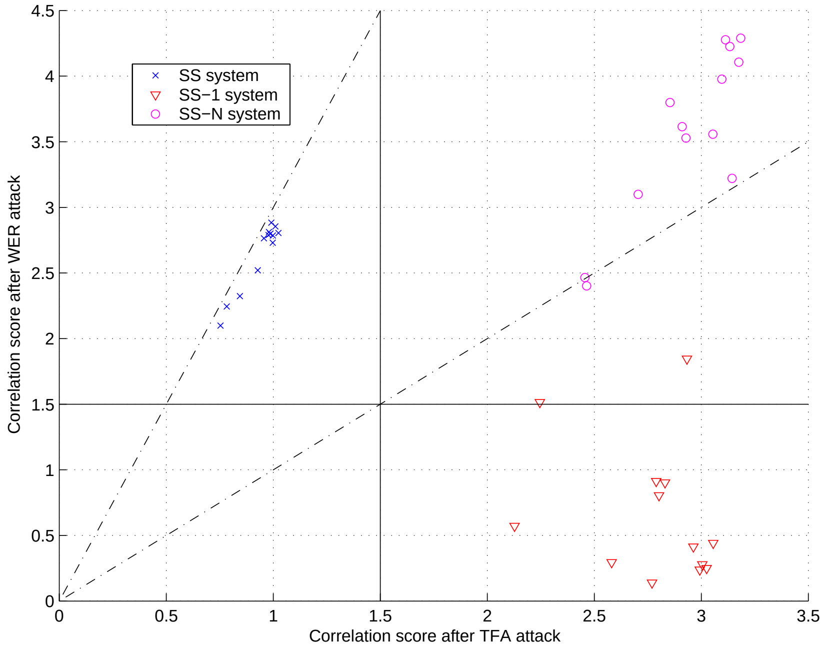

Figure 4.2: Resilience of the three presented watermarking systems (SS, SS-1 and SS-N) against intra-video collusion attacks.

attack on the other hand. Finally, the correlation score is computed for all the videos. The obtained results have been gathered in Figure 4.2.

Each watermarking scheme is represented by a specific symbol: crosses for video shots watermarked with SS system, triangles for those with SS-1 system and circles for those with SS-N system. The figure has also been divided into four quadrants whose borders are defined by the detection threshold $\tau_{detect} = 1,5$. The very first observation is that each kind of point is mainly located in a single quadrant. The crosses of the SS system are located in the upper-left quadrant, which confirms that this system resists the watermark estimation remodulation attack while it is weak against temporal frame averaging. In fact, theoretical results have shown that the circles should have coordinates like $(x/w, x)$. One can easily check that the circles are in the neighborhood of the line defined by $y = wx$, which has been drawn on the figure for the window size $w = 3$ used in the experiments. On the other hand, the triangles of the SS-1 system are in the lower-right quadrant, which supports the theoretical results asserting that this system is robust against temporal frame averaging while the watermark estimation remodulation attack succeeds in stirring out the watermarked signal. In fact, one can also notice that this second attack is more or less efficient depending on the video content of the shot. If a static shot is considered, the estimation is poorer and the attack less effective i.e. the associated triangle lies in the upper-right quadrant. Being in this quadrant means that the video shot still contains a watermark either after temporal frame averaging, or the watermark estimation remodulation attack. As a result, finding all the circles of the SS-N system in this region of the figure confirms the theoretical prediction that the algorithm resists both attacks. The watermark estimation remodulation attack even increases the correlation score as asserted in Equation 4.10.

Chapter 5

Attacker Options

It has been previously shown that the introduced SS-N system is robust against two formerly designed intra-video collusion attacks. However, the attackers are likely to modify and adjust their approach according to the newly released watermarking scheme. Obviously, it is useless to perform an attack based on temporal frame averaging since the detector has been designed to detect watermarks whose energy has been potentially spread between several video frames. There is also no reason to launch a watermark estimation procedure, which outputs a single watermark, if it is known that several alternative watermarks have been embedded. Basically, the SS-N system relies on the secrecy of the set of watermarks $\{W_i\}$. If this set is disclosed, the attacker can find out which watermark has been embedded in each frame and perform a simple remodulation to remove it. As a result, the next subsections will verify that the set of secret watermarks cannot be easily estimated in a blind manner.

5.1 Brute Force Approach

According to Kerckhoffs' principle [17], the watermarking system is publicly known and the attacker is aware that N alternative watermarks have been randomly embedded in the video. One approach consists then in distributing the watermarked video frames between N sets $\mathcal{S}_i$ and estimating the watermarks $\tilde{W}_i$ from each one of those sets. Assuming that the attacker has access to the ideal watermark estimator $\mathcal{E}_0(.)$, the following watermarks are obtained:

$$\tilde{W}_i = \sum_{j=1}^{N} n_{i,j} W_j \tag{5.1}$$

where $n_{i,j}$ is the number of frames carrying the watermark W_j in the set $\mathcal{S}_i$. Moreover, if each set $\mathcal{S}_i$ contains P frames, the $n_{i,j}$'s verify:

$$\sum_{i=1}^{N} n_{i,j} = P \quad \text{and} \quad \sum_{j=1}^{N} n_{i,j} = P \tag{5.2}$$

Once those watermark estimations have been obtained, the attacker should have a criterion to determine if one or more watermarks have been correctly estimated.

When the N sets $\mathcal{S}_i$ are built, the $n_{i,j}$ are unknown. The attacker can only compute the different correlations between the estimated watermarks $\{\tilde{W}_i\}$ defined as follows:

$$c_{i_1,i_2} = \tilde{W}_{i_1} \cdot \tilde{W}_{i_2} = \sum_{j=1}^{N} n_{i_1,j}.n_{i_2,j} \tag{5.3}$$

For a given estimated watermark $\tilde{W}_{i_0}$, the sum of the correlations with the set of estimated watermarks $\{\tilde{W}_i\}$ is equal to:

$$\sum_{i=1}^{N} c_{i_0,i} = \sum_{i=1}^{N}\sum_{j=1}^{N} n_{i_0,j}.n_{i,j} = P^2 \tag{5.4}$$

Now, let assume that there exists an index i_0 for which $c_{i_0,i_0} > mP^2$ with m in $[0.5, 1]$. It can be shown that $n_{i_0,j^*} = \max_j(n_{i_0,j})$ is greater than mP. Since m is greater than 0.5, the correlation score between $\tilde{W}_{i_0}$ and the video frames carrying W_{j^*} is higher than with the other ones. As a result, the attacker can distinguish the frames carrying W_{j^*}, obtain a finer estimate for W_{j^*} and iterate the attack with a reduced set of video frames to estimate the remaining watermarks i.e. with a reduced complexity.

In summary, this demonstrates that the attacker can remove the embedded watermark. However, the complexity of this brute force approach is very high. Since the probability that $c_{i,i}$ is greater than mP^2 is difficult to estimate, the probability that $n_{i,j}$ is greater than mP will be considered below to obtain a lower bound for the complexity. It is quite straightforward that the probability that $n_{i,j}$ is equal to n is given by:

$$\begin{aligned} P\Big(n_{i,j} = n\Big) &= \binom{P}{k} P\Big(W = W_i\Big)^k P\Big(W \neq W_i\Big)^{P-k} \\ &= \binom{P}{k}\Big(\frac{1}{N}\Big)^k\Big(1 - \frac{1}{N}\Big)^{P-k} \end{aligned} \tag{5.5}$$

Those probabilities can then be summed to obtain the probability p_L that $n_{i,j}$ is strictly greater than L.

$$p_L = \sum_{n=\lfloor L+1 \rfloor}^{P} \binom{P}{k}\Big(\frac{1}{N}\Big)^k\Big(1 - \frac{1}{N}\Big)^{P-k} \tag{5.6}$$

Proposition 5 *The attacker is required to compute at least $O(N^{\lfloor mP+4 \rfloor})$ linear correlations between estimated watermarks to terminate this brute force attack.*

Proof When N grows large, the probability p_L is almost reduced to a single term:

$$\begin{aligned} p_L &\approx \binom{P}{\lfloor L+1 \rfloor}\Big(\frac{1}{N}\Big)^{\lfloor L+1 \rfloor}\Big(1 - \frac{1}{N}\Big)^{P-\lfloor L+1 \rfloor} \\ &\approx \binom{P}{\lfloor L+1 \rfloor} N^{-\lfloor L+1 \rfloor} \end{aligned} \tag{5.7}$$

As a result, the attacker should in average distribute the video frames between N sets $1/p_L$ times before obtaining a distribution that can be exploited. Moreover, for each one of those distributions, $N(N+1)/2$ correlations between estimated marks are computed. In other terms, the number of correlation n_{corr}^N is equal to:

$$\begin{aligned} n_{corr}^N &= \frac{1}{2}\binom{P}{\lfloor L+1 \rfloor}^{-1} N^{\lfloor L+1 \rfloor} N(N+1) \\ &\approx \frac{1}{2}\binom{P}{\lfloor L+1 \rfloor}^{-1} N^{\lfloor L+3 \rfloor} \end{aligned} \tag{5.8}$$

When the attacker has performed all those operations, a single watermark has been estimated and the process has to be continued to estimate the $N-1$ remaining watermarks. Consequently the total number of computed correlations is equal to:

$$n_{corr} = \sum_{i=2}^{N} n_{corr}^i \approx \frac{1}{2\lfloor L+4 \rfloor}\binom{P}{\lfloor L+1 \rfloor}^{-1} N^{\lfloor L+4 \rfloor} \tag{5.9}$$

In practice, L is equal to mP with m in $[0.5, 1]$ and the asserted result is obtained. $\blacksquare$

Checking for the presence of N alternative watermarks in each video frame increases the complexity of the algorithm by a factor N. On the attacker side, the complexity of a brute force attack is proportional to $N^{\lfloor mP+4 \rfloor}$. For example, for $N = 64$, $P = 50$ and $m = 0.5$, Equation 5.9 means that the brute force attack requires at least 3.10^{36} correlation computations, which makes the attack unpractical.

5.2 Vector Quantization

The brute force attack basically exploits a previously designed attack, which is watermark estimation remodulation. However, it might be more efficient to design a new attack, specifically adjusted to the new watermarking algorithm. In particular, the several individual watermark estimates obtained, from each video frame, can be regarded as several vectors $\{v_t\}$ in a very high dimensional space. Since those vectors should approximate the embedded watermarks $\{W_i\}$, the whole problem comes down to vector quantization [11]. In other terms, it should be possible to define N clusters $\mathcal{C}_i$ whose centroids c_i are good estimates of the embedded watermarks. When a single watermark has been embedded in all the video frames (SS-1 system), this approach is completely equivalent to the watermark estimation remodulation attack: the single centroid is set equal to the average of all the vectors i.e. watermark estimates. On the other hand, when the same watermark is never inserted twice (SS system), the clusters will contain a single watermark estimate, which is a poor estimate of the embedded watermark, and the attack is a failure. In summary, the vector quantization approach only makes sense if enough individual estimates are available for each watermark to be estimated.

The k-means algorithm is a very simple way to perform vector quantization. In a first step, the vectors $\{v_t\}$ are distributed among the different clusters $\{\mathcal{C}_i\}$, so that each vector is assigned to the cluster associated with its nearest centroid c_i according to a given distance e.g. the Euclidean distance. In a second step, the centroids are updated and the algorithm iterates until convergence. Additionally, this original algorithm has been slightly modified so that non-pertinent estimated watermark samples, i.e. the samples whose value is greater than τ_{estim} in absolute value, are discarded during distance and centroid computations. This algorithm has been proven to converge toward a position which locally minimizes the sum of the distance from the estimated watermarks to their associated centroid. However, nothing insures that the *perfect* estimation, i.e. when the centroids are equal to $\{W_i\}$, is a local minimum. In fact, experiments have shown that the algorithm iterates for a while before terminating when the set of embedded watermarks $\{W_i\}$ is given as the initial guess. In other terms, the algorithm makes the centroids move a bit away, which means that the optimal estimation will never be reached after convergence, whatever the initial guess is. Moreover, the attacker has in practice no a priori knowledge of the centroids and should use a random initial guess, which has an impact on the finally reached local minimum.

Experiments have shown that this naive algorithm does not usually succeed in converging toward a fine enough estimation of the watermarks $\{W_i\}$ to trap the detector. Nevertheless, this algorithm can be further improved. First, the k-means algorithm can be regarded as a simplified version of the more general Expectation-Maximization (EM) algorithm for Gaussian Mixture Models with equal variances for the Gaussians and hard decisions i.e. each vector is assigned to single cluster at each iteration. Thus, it can be interesting to consider a general Gaussian mixture and to perform the EM algorithm with soft decisions [22] i.e. each vector is assigned to each cluster with a given probability. The second potential improvement is related to the initial guess. Randomness makes indeed the final position not reliable and a splitting strategy as in [21] might be more efficient. Initially, the centroid c_0 of all the vectors is split into two centroids $c_0 \pm \epsilon$ where ϵ is a small perturbation along the direction of principal variation. The traditional k-means algorithm iterates then with two centroids. When convergence is reached, each centroid is split, iterations are made until convergence and so on until the desired number of centroids is obtained. The last improvement is related with the fact that incomplete data are used i.e. some samples of the vectors $\{v_t\}$ are not pertinently estimated and are ignored. Instead of discarding such samples, an alternative consists in completing the missing data with a maximum likelihood criterion [7]. Those improvements may make the vector quantization approach a success but it cannot be denied that the theoretical requirements to perform such a successful attack have been raised. With the former watermark estimation remodulation attack, image filtering and temporal averaging were indeed the only requirements.

Chapter 6

Conclusion

To date, very few watermarking schemes (SLIDE algorithm [26]) have addressed the intra-video collusion issue. Most of the existing systems relies indeed on a frame-by-frame approach [15, 23], so that previous results obtained for still images can be exploited. Unfortunately, this results in weak algorithms in terms of security. Thus, a new watermarking system (SS-N system) based on the spread spectrum theory has been introduced in this paper, which can be seen as some generalization of existing systems (SS and SS-1 systems). The embedding process basically consists in inserting a watermark randomly chosen from a finite set of key-dependent pseudo-random watermarks $\{W_i\}$. On the other side, the detector blindly checks for the presence of all the watermarks $\{W_i\}$ in each incoming video frame and has the asset not to require any temporal synchronization. Furthermore, the system has been proven to be robust against two common intra-video collusion attacks, which defeat the previous watermarking systems. The only noticeable drawback of the proposed system is that checking for the presence of several watermarks in each video frame may prevent real-time detection. On the attacker side, a brute force approach has been shown to be unpractical to estimate the set of watermarks $\{W_i\}$ and a potentially successful attack based on vector quantization has been proposed. However, such an attack requires a much higher level of expertise in signal processing than previously. Further efforts have also to be made to resist more advanced intra-video collusion attacks [9].

If the introduced watermarking system has improved performances in terms of security, several tracks remain for future research to obtain a complete secure watermarking system:

Capacity Currently, the SS-N system has no capacity. It only checks for the presence of a set of watermarks. However, some applications require that a watermark encoding a multi-bit message is embedded. For example, in a video-on-demand framework, the embedded watermark should encode the identity of the customer so that he/she can be identified if an illegal copy of the video is found. The codewords need to be chosen cautiously [1] so that inter-video collusion is not possible.

Security Security against collusion has been enhanced with the proposed approach in comparison with previous systems. However, it has also been demonstrated that the scheme can still be broken with a higher level of expertise. Embedding either one watermark or another one allows indeed attacks based on vector quantization. It should be interesting to investigate how the N alternative watermarks can be randomly mixed in each video frame so that $1/N$ percent of the resulting watermark samples are from a given watermark of the set $\{W_i\}$.

Visibility Embedding alternative watermarks in successive video frames is likely to introduce some annoying visible artifacts due to flicker noise [30]. As a result, frame hashing [10] might be introduced to monitor the changes of the watermark and subsequently control the flicker noise.

Spatial synchronization The SS-N system does not require any temporal synchronization. However, spatial desynchronization (cropping, rotation, translation, scaling) is likely to defeat the system. A common counterattack consists in introducing a fixed known template [18] so that the transformation can be estimated and inverted. Unfortunately, such a template is not secure [13] and becomes the Achilles heel of the system. As a result, an alternative solution has to be found to be able to face spatial desynchronization.

Watermarking compressed data All the presented systems embed a secret watermark in uncompressed video frames. Since video data is often compressed for storage and/or transmission, one should investigate how the embedding can be made directly in the compressed domain.

Bibliography

[1] D. Boneh and J. Shaw. Collusion secure fingerprinting for digital data. *IEEE Transaction on Information Theory*, 44(5):1897–1905, September 1998.

[2] Certimark: http://www.certimark.org.

[3] H. Cohen. *A Course in Computational Algebraic Number Theory*. Springer-Verlag, 1993.

[4] I. Cox and M. Miller. Preprocessing media to facilitate later insertion of a watermark. In *Proceedings of the International Conference on Digital Signal Processing*, volume 1, pages 67–70, July 2002.

[5] I. Cox, M. Miller, and J. Bloom. *Digital Watermarking*. Morgan Kaufmann Publishers, 2001.

[6] F. Deguillaume, G. Csurka, J. Ó Ruanaidh, and T. Pun. Robust 3D DFT video watermarking. In *Proceedings of SPIE 3657, Security and Watermarking of Multimedia Contents*, pages 113–124, January 1999.

[7] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.

[8] G. Doërr and J.-L. Dugelay. A guide tour of video watermarking. *Signal Processing: Image Communication*, 18(4):263–282, April 2003.

[9] G. Doërr and J.-L. Dugelay. New intra-video collusion attack using mosaicing. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, July 2003.

[10] J. Fridrich and M. Goljan. Robust hash functions for digital watermarking. In *Proceedings of the International Conference on Information Technology: Coding and Computing*, pages 178–183, March 2000.

[11] A. Gersho and R. Gray. *Vector Quantization and Signal compression*. Kluwer Academic Publishers, 1992.

[12] F. Hartung and B. Girod. Watermarking of uncompressed and compressed video. *Signal Processing*, 66(3):283–301, May 1998.

[13] A. Herrigel, S. Voloshynovskiy, and Y. Rytsar. The watermark template attack. In *Proceedings of SPIE 4314, Security and Watermarking of Multimedia Contents III*, pages 394–405, January 2001.

[14] F. Jordan, M. Kutter, and T. Ebrahimi. Proposal of a watermarking technique for hiding/retrieving data in compressed and decompressed video. In *JTC1/SC29/WG11 MPEG97/M2281*. ISO/IEC, July 1997.

[15] T. Kalker, G. Depovere, J. Haitsma, and M. Maes. A video watermarking system for broadcast monitoring. In *Proceedings of SPIE 3657, Security and Watermarking of Multimedia Contents*, pages 103–112, January 1999.

[16] S. Katzenbeisser and F. Petitcolas. *Information Hiding: Techniques for Steganography and Digital Watermarking*. Artech House, 1999.

[17] A. Kerckhoffs. La cryptographie militaire. *Journal des sciences militaires*, IX:5–83, January 1883.

[18] M. Kutter. Watermarking resisting to translation, rotation and scaling. In *Proceedings of SPIE 3528, Multimedia Systems and Applications*, pages 423–431, November 1998.

[19] M. Kutter and F. Petitcolas. A fair benchmark for image watermarking systems. In *Proceedings of SPIE 3657, Security and Watermarking of Multimedia Contents*, pages 226–239, January 1999.

[20] G. Langelaar, R. Lagendijk, and J. Biemond. Real-time labelling of MPEG-2 compressed video. *Journal of Visual Communication and Image Representation*, 9(4):256–270, August 1998.

[21] Y. Linde, A. Buzo, and R. Gray. An algorithm for vector quantizer design. *IEEE Transactions on Communications*, 28(1):84–95, January 1980.

[22] G. McLachlan and D. Peel. *Finite Mixture Models*. Wiley-Interscience, 2000.

[23] B. Mobasseri. Exploring CDMA for watermarking of digital video. In *Proceedings of SPIE 3657, Security and Watermarking of Multimedia Contents*, pages 96–102, January 1999.

[24] B. Mobasseri, M. Sieffert, and R. Simard. Content authentication and tamper detection in video. In *Proceedings of the IEEE International Conference on Image Processing*, volume I, pages 458–461, September 2000.

[25] R. Pickholtz, D. Schilling, and L. Millstein. Theory of spread spectrum communications - a tutorial. *IEEE Transactions on Communications*, 30(5):855–884, May 1982.

[26] K. Su, D. Kundur, and D. Hatzinakos. A novel approach to collusion resistant video watermarking. In *Proceedings of SPIE 4675, Security and Watermarking of Multimedia Contents IV*, pages 491–502, January 2002.

[27] M. Swanson, B. Zhu, and A. Tewfik. Multiresolution scene-based video watermarking using perceptual models. *IEEE Journal on Selected Areas in Communications*, 16(4):540–550, May 1998.

[28] C. Voloshynovskiy, S. Pereira, A. Herrigel, N. Baumgärtner, and T. Pun. Generalized watermarking attack based on watermark estimation and perceptual remodulation. In *Proceedings of SPIE 3971, Security and Watermarking of Multimedia Contents II*, pages 358–370, January 2000.

[29] S. Voloshynovskiy, A. Herrigel, N. Baumgärtner, and T. Pun. A stochastic approach to content adaptive digital image watermarking. In *Proceedings of the Third International Workshop on Information Hiding (LNCS 1768)*, pages 211–236, September 1999.

[30] S. Winkler, E. Gelasca, and T. Ebrahimi. Perceptual quality assessment for video watermarking. In *Proceedings of the International Conference on Information Technology: Coding and Computing*, pages 90–94, April 2002.