Optimal detection of multiplicative watermarks

Job Oostveen, Ton Kalker, Jean-Paul Linnartz*

ABSTRACT

We derive a watermark detector for images which are watermarked in a multiplicative way. Under the assumptions that the watermark coefficients are a known, binary valued sequence and that the original image coefficients are an i.i.d. random sequence from a Weibull distribution, we show that this watermark should be detected by raising the observations to the power β before correlating them with the watermark (here β is the parameter in the exponent of the Weibull probability density function). The approach is based on maximum-likelihood estimation of the embedding strength of the watermark. The result is illustrated by experiments and an extension to Gaussian distributed data is discussed.

1 Introduction

An important issue in digital watermarking is the conflict between robust detection and perceptibility. On the one hand, a watermarked image should only be an imperceptible modification of the original image, but at the same time the watermark should be robustly detectable, which is easiest if embedding the watermark leads to a considerable change of the image. To be able to have practical watermarking schemes, it is of eminent importance to have good detection schemes available.

In the literature it is very common that the watermark is embedded in an additive way, i.e.,

$$q_i \quad = \quad p_i + sw_i,$$

where $\{w_i\}$ is the watermark sequence, $\{p_i\}$ is the original image, $\{q_i\}$ is the watermarked image and s is the strength of embedding. In the Gaussian case, if the image $\{p_i\}$ is unknown, such an additive watermark is optimally detected by correlation with the watermark. This can be proven using a version of the matched filtering theorem. Many refinements and improvements of this detection scheme are known, for instance by using whitening or Wiener filters. We cannot give an exhaustive list of articles dealing with optimal detection of additive watermarks. Instead, we refer to the very recent papers [3], [4], [5], [6].

In this paper we want to look at a different way of embedding watermarks: multiplicative embedding. The watermarked coefficients q_i are now formed from the watermark coefficients w_i and the original image coefficients p_i according to

$$q_i \quad = \quad p_i(1 + sw_i), \tag{1}$$

where s is the embedding strength. This way of embedding was proposed, among others, by Cox et.al. [1]. It provides a way of perceptual masking of the watermark in the image. It is well-known that straightforwardly embedding a watermark in an additive way will result in an image with perceptible artefacts. The perceptual effect of the watermark can be decreased by using perceptual masking, for instance Weber's law. It tells us that for images the luminance of a pixel is a useful perceptual mask. This mask would lead to the multiplicative embedding studied in this paper, as in equation (1).

It is unlikely that straightforward correlation with the watermark pattern would be the optimal manner of detecting the presence of the watermark in the multiplicative case, as well. Optimal detection, in the sense of Bayes criterion or the Neyman-Pearson Criterion, leads to a likelihood ratio test: hypothesis H_0 is accepted if the likelihood ratio

$$\Lambda(Q) \quad = \quad \frac{f_Q(Q|H_0)}{f_Q(Q|H_1)}$$

exceeds a certain threshold, where Q is the observation and f_Q is the probability density function of Q. In the case of a simple hypothesis H_0 and alternative H_1 this leads to a tractable problem. In our case, however, the hypothesis is simple ($H_0 : \ s = 0$) and the alternative is composite ($H_1 : \ s \neq 0$). The composite alternative leads to a complication in computing the likelihood ratio. We have

$$f_Q(Q|H_1) \quad = \quad \int f_Q(Q|s) f_s(s|H_1) ds,$$

[1] Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands. E-mail: job.oostveen@philips.com, ton.kalker@ieee.org, j.p.linnartz@philips.com

and so we will have to know the probability distribution of s given H_1. We do not want to assume this knowledge, and so the likelihood ratio test becomes unusable for our problem. Therefore we take a different approach, based on *maximum-likelihood estimation* of s.

De Rosa et.al [2] consider optimal detection for the same type of multiplicative embedding. They derive an optimal detector based on a likelihood ratio test for image coefficients satisfying a Weibull distribution. They circumvent the problem sketched above by taking a fixed value of s, i.e., by testing $H_0 : s = 0$ against $H_1 : s = s^*$, for a fixed value s^*. For this it would be necessary to have a good guess for the actual value of s. Of course, knowledge of the embedding process could give some information about s. Still we think it is not reasonable to assume knowledge on the actual value of s, since the effect of many attacks on the watermarked image can roughly be modelled by a decrease of this value.

Note that for the present discussion of detection methods it is not relevant whether p_i and q_i are spatial variables (like the luminance values of a picture), temporal variables (like the sound intensity in an audio frame), frequency variables (like DFT or DCT coefficients) or any other set of representative variables. Of course, the particular choice may have a strong impact on perceptibility and robustness of the watermark, as well as on which statistical model is suitable for the image coefficients.

2 Main result

In the present section we formulate and prove our main result: the derivation of a detector for multiplicative watermarks based on maximum-likelihood estimation of s. This is done under the assumption that the original data are modelled by a Weibull probability density function with parameters α and β:

$$f(x) \quad = \quad \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, \qquad (2)$$

We show that under this assumption, our detector consists of raising the data to the power β and subsequently correlating with the watermark. We do not restrict a priori what values s can take, except that s should be sufficiently small, so that the watermark is imperceptible.

DCT and DFT coefficients are usually modelled by a Weibull distribution (See for instance De Rosa et.al [2]). The Rayleigh distribution is a special case, corresponding to $\beta = 2$. We also formulate our result for the Gaussian case, because the analysis turns out to be completely similar to the case of the Weibull distribution

In the sequel, we use the following shorthand notation: $\langle a, b\rangle = \frac{1}{N}\sum_{i=1}^{N} a_i b_i$ and $\|a\| = \sqrt{\langle a, a\rangle}$.

Theorem 2.1 *Consider the sequences $P = \{p_i : i = 1, \ldots, N\}$, $W = \{w_i : i = 1, \ldots, N\}$ and $Q = \{q_i : i = 1, \ldots, N\}$. Assume that P is an i.i.d. sequence of stochastic variables, drawn from a Weibull distribution with parameters α and β (see eqn. (2)). Moreover assume that W is a known, zero-mean, binary valued sequence (i.e., $\forall i : w_i = \pm 1$).*

The maximum-likelihood decision variable is given by

$$d \quad = \quad \frac{1}{N}\sum_{i=1}^{N} q_i^{\beta} w_i. \qquad (3)$$

The idea of the proof is as follows. We derive a maximum-likelihood estimator for s. That is, we derive the joint probability density function $f(Q; s)$ for the observations $Q = \{q_i\}$, given s. The maximum-likelihood estimate $\hat{s}$ is the value of s which maximises f for the observed values of q_i. It follows that $\hat{s}$ is proportional $d = \frac{1}{N}\sum_{i=1}^{N} q_i^{\beta} w_i$.

The requirement that the watermark be zero-mean comes down to partitioning the index set $\{1, \ldots, N\}$ into two subsets, one of which corresponds to $w_i = 1$ and the other to $w_i = -1$. The requirement that p_i are identically distributed can be achieved by pre-whitening the data. How exactly this should be done in the context of multiplicative watermarks is the subject of further research.

To be able to set a threshold for the detection, we need to have some information about the probability distribution of d as a function of s.

Theorem 2.2 *Consider the situation of Theorem 2.1. d satisfies*

$$\begin{aligned} E[d|s, W] &= s\beta\alpha^{\beta} + \mathcal{O}(s^3), \\ \operatorname{var}(d|s, W) &= \frac{1}{N}(1 + (2\beta^2 - \beta)s^2)\alpha^{2\beta} + \mathcal{O}(s^3). \end{aligned}$$

An important measure of the strength of an estimation method is the quotient between the expected value of d (depending on s) and the standard deviation of d for $s = 0$. In the case of Rayleigh distributed data and the detection method of Theorem 2.1, we obtain

$$\frac{E[d|s, W]}{\sqrt{\operatorname{var}(d|s = 0, W)}} \quad = \quad 2\sqrt{N}s.$$

If, instead, we would use linear correlation (i.e., $d_l = \langle Q, W\rangle$), we have

$$E[d_l|s, W] \; = \; s, \qquad \operatorname{var}(d_l|s, W) \; = \; \frac{\sigma^2(4 - \pi)}{2N},$$

and so the quotient would be

$$\frac{E[d_l|s, W]}{\sqrt{\operatorname{var}(d_l|s = 0, W)}} \quad = \quad \frac{\sqrt{2}\sqrt{N}s}{\sigma\sqrt{4 - \pi}},$$

which differs from the value for correlation with squared observation by a factor $\frac{\sigma\sqrt{4-\pi}}{\sqrt{2}}$.

3 Extension to the Gaussian distribution

The results in the previous section extend straightforwardly to the case of p_i having a zero-mean Gaussian distribution with variance σ^2. We will not repeat proofs, but just formulate the result

Theorem 3.1 *Let p_i be an i.i.d. sequence drawn from a zero-mean Gaussian distribution with variance σ^2. Moreover, assume that w_i is a binary valued, zero-mean sequence, and let $q_i = p_i(1 + sw_i)$. The maximum-likelihood detector is given by (3).*

Furthermore, d satisfies

$$\begin{aligned} E[d|s, W] &= 2s\sigma^2 \\ \mathrm{var}(d|s, W) &= 2\sigma^4(1 + 6s^2 + s^4)/N. \end{aligned}$$

4 Experimental results

The theory of the previous sections has been applied to the case of audio watermarking by means of modulation of Fourier coefficients. For the experiment a well known 15 second audio clip x (*donna*) is watermarked by convolving with a length 1023, DC-free, zero-phase and normally distributed random filter w. After scaling the convolved sequence y to have the same standard deviation as x, the watermarked clip is obtained as $z = x + \alpha y$, where α controls the energy of the watermark. It is not difficult to see that the described procedure is equivalent to multiplicative embedding in the frequency domain, modulating with the Fourier transform W of w.

For detection, (a selected interval of) the watermarked clip z is convolved with a simple high pass FIR filter, *cyclically folded* to a length 1024 sequence, and transformed to a spectral representation Z by means of a fast Fourier transform. Fitting the sequence $|Z|$ to a Weibull distribution, we find a Weibull exponent $\beta = 0.95$ Raising $|Z|$ to the power γ, $\gamma = 0.5, \ldots, 2$ and computing the inner product with W, a decision value $d'(\alpha, \gamma)$ is obtained. By computing the ratio of this number with the standard deviation of a large number of detection results obtained with false watermarks W', a reliability measure $d(\alpha, \gamma)$ is obtained. The results of the experiments are presented in Figure 1.

The figure clearly shows that optimal reliability is achieved for a value of γ slightly smaller than 1, in accordance with the value of the Weibull exponent β.

5 Conclusions

In this paper we have derived an optimal detector for multiplicative watermark, under the assumption that the image coefficients are distributed according to a Weibull distribution with parameters α and β. Our derivation shows that the observations should be raised to the power β before correlation. The resulting detector differs from those used nowadays, which are based on correlating the observations with the watermark. This result shows that very likely it is possible to improve on the common practice to use correlation of a watermark directly with the data to detect additive watermarks which are embedded using perceptual masks (the result of which embedding can be modelled by a multiplicative watermark).

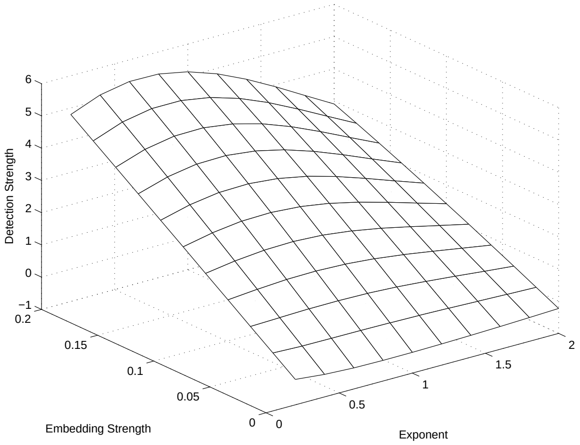

Figure 1: Detection reliability for varying embedding strength and detection exponent.

6 Proofs

In this section we give the proofs of our theorems

Proof of Theorem 2.1 As p_i is distributed according to a Weibull distribution with parameters α and β, the joint probability density function of $\{q_i\}$, where $q_i = p_i(1 + sw_i)$ and s is viewed as a parameter, is given by

$$f(Q; s) = \prod_{i=1}^{N} \frac{\beta q_i}{\alpha^2(1 + sw_i)^2} e^{-\left(\frac{q_i}{\alpha(1+sw_i)}\right)^{\beta}}.$$

The paradigm of maximum-likelihood estimation now means that we estimate s such that the estimate $\hat{s}$ maximises the above joint probability density. It is equivalent, but notationally easier, to maximise the logarithm of f:

$$\begin{aligned} L(Q; s) &= N\log(\beta) - 2N\log(\alpha) - 2\sum_{i=1}^{N}\log(1 + sw_i) \\ &\quad + \sum_{i=1}^{N}\log(q_i) - \sum_{i=1}^{N}\left(\frac{q_i}{\alpha(1 + sw_i)}\right)^{\beta}. \end{aligned}$$

Now, we need to solve

$$\begin{aligned} \frac{\partial L}{\partial s} &= -2\sum_{i=1}^{N}\frac{w_i}{1 + sw_i} + \sum_{i-1}^{N}\frac{\beta q_i^{\beta} w_i}{\alpha^{\beta}(1 + sw_i)^{\beta+1}} \\ &= 0. \end{aligned}$$

Because of imperceptibility requirements, it is reasonable to assume that $|s|$ is small. Therefore, we replace

$\partial L/\partial s$ by its first order Taylor expansion

$$\begin{aligned}\frac{\partial L}{\partial s} &\approx \sum_{i=1}^{N} -2w_i(1-sw_i) + \frac{\beta q_i^\beta w_i}{\alpha^\beta}(1-(\beta+1)sw_i) \\ &= 0.\end{aligned}$$

Solving for s and using the fact that $\{w_i\}$ is a zero-mean sequence leads to

$$\hat{s} = \frac{\beta \frac{1}{N}\sum_{i=1}^{N} q_i^\beta w_i}{(\beta^2+\beta)\frac{1}{N}\sum_{i=1}^{N} q_i^\beta w_i^2 - 2\alpha^\beta w_i^2}.$$

Using the fact that $w_i^2 = 1$, we obtain

$$\hat{s} = \frac{\beta\langle q_i^\beta, w_i\rangle}{(\beta^2+\beta)\|Q^\beta\| - 2\alpha^\beta}.$$

■

Proof of Theorem 2.2 This proof is a matter of long but straightforward computations. All summations are over the range $1,\ldots,N$. Using some standard integration tricks, it can be computed that $Ep_i^\beta = \alpha^\beta$, $Ep_i^\beta p_j^\beta = (1+\delta_{ij})\alpha^{2\beta}$.

First,

$$\begin{aligned}&E[d|s,W] \\ &\quad= E[\frac{1}{N}\sum_i q_i^\beta w_i\ |s,W] = \frac{1}{N}\sum_i (1+sw_i)^\beta w_i E p_i^\beta \\ &\quad= \frac{\alpha^\beta}{N}\sum_i \left(w_i + \beta s w_i^2 + \frac{1}{2}(\beta^2-\beta)s^2 w_i^3\right) + \mathcal{O}(s^3) \\ &\quad= \frac{\alpha^\beta \beta s}{N}\sum_i w_i^2 = s\beta\alpha^\beta.\end{aligned}$$

Secondly,

$$\begin{aligned}&E[d^2|s,W] \\ &\quad= \frac{1}{N^2}\sum_{i,j}(1+sw_i)^\beta(1+sw_j)^\beta w_i w_j E p_i^\beta p_j^\beta \\ &\quad= \frac{1}{N^2}\sum_{i,j}(w_i + \beta s w_i^2 + \frac{1}{2}(\beta^2-\beta)s^2 w_i^3)\cdot \\ &\qquad\quad (w_j + \beta s w_j^2 + \frac{1}{2}(\beta^2-\beta)s^2 w_j^3)E p_i^2 p_j^2 + \mathcal{O}(s^3) \\ &\quad= \frac{a^\beta}{N^2}\sum_i w_i^2 + \beta^2 s^2 w_i^4 + 2\beta s w_i^3 + (\beta^2-\beta)s^2 w_i^4 \\ &\qquad + \frac{a^\beta}{N^2}\left(\sum_i \beta s w_i^2\right)^2 + \mathcal{O}(s^3) \\ &\quad= \alpha^{2\beta}\beta^2 s^2 + \frac{\alpha^{2\beta}}{N}(1+(2\beta^2-\beta)s^2) + \mathcal{O}(s^3),\end{aligned}$$

where we used the fact that summations over odd powers of w_i or w_j are equal to zero. Using this, we obtain

$$\begin{aligned}\mathrm{var}(d|s,W) &= E[d^2|s,W] - (E[d|s,W])^2 \\ &= \frac{1}{N}(1+s^2(2\beta^2-\beta))\alpha^{2\beta} + \mathcal{O}(s^3).\end{aligned}$$

■

References

[1] I.J. Cox, J. Killian, F. Thomson Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6:1673–1687, 1997.

[2] A. de Rosa, M. Barni, F. Bartolini, V. Cappellini, and A. Piva. Optimum decoding of non-additive full frame DFT watermarks. In *Proceedings of the Third International Information Hiding Workshop*, pages 167–179, Dresden, 1999.

[3] J.R. Hernandez and F. Perez-Gonzalez. Statistical analysis of watermarking schemes for copyright protection of images. *Proceedings of the IEEE*, 87:1142–1166, 1999.

[4] J.P. Linnartz, G. Depovere, and T. Kalker. On the design of a watermarking system: considerations and rationales. In *Proceedings of the Third International Information Hiding Workshop*, pages 303–314, Dresden, 1999.

[5] M.L. Miller and J.A. Bloom. Computing the probability of false watermark detection. In *Proceedings of the Third International Information Hiding Workshop*, pages 154–166, Dresden, 1999.

[6] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun. A stochastic approach to content adaptive digital image watermarking. In *Proceedings of the Third International Information Hiding Workshop*, pages 219–244, Dresden, 1999.