MAXIMUM LIKELIHOOD DETECTION OF MULTIPLICATIVE WATERMARKS

Ton Kalker, Job Oostveen, Jean-Paul Linnartz*

Consider a watermarked image, in which the watermark is embedded multiplicatively (i.e, $q_i = p_i(1+w_i)$*). Under the assumption that the watermark is a known, binary valued sequence and the original image is an i.i.d. random sequence drawn from a Rayleigh distribution, we show that this watermark should be detected by squaring the observations before correlating them with the watermark. This is proven using a maximum likelihood estimation approach. Extensions to the Gaussian and Weibull distribution are possible.*

INTRODUCTION

Recently, watermarking of audio and video material has received much attention. This interest has been driven by the fast digitalisation of such content. In their digital form, audio and video (or content, for short) can easily be stored and transmitted, but the downside is that they can be easily copied, as well, without loss of quality. To enable trade of electronic media, it is therefore of eminent importance that new techniques for copyright protection are developed.

Digital watermarking forms an important building block of copyright protection schemes. A digital watermark is an additional piece of information which is embedded in the content. The following is a list of the most common requirements for watermarking: the watermark should

1. be imperceptible;
2. survive A/D and D/A conversion and common compression techniques;
3. be robust against geometric distortions (e.g., scaling).

The imperceptibility requirement is, among others, a restriction on the allowable energy of the watermark, relative to the energy of the content. The requirements 2 and 3, above, not only mean that the watermark should be left intact by the mentioned operations, but also that a distorted watermark should still be detectable. All of this shows the necessity of having a very strong and reliable detection method.

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

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

$$q_i = p_i + s w_i,$$

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. 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

$$q_i = p_i(1 + s w_i), \tag{1}$$

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. Weber's law tells us that for images the luminance of a pixel is a useful perceptual mask. It is well-known that straightforwardly embedding a watermark in an additive way will result in an image with perceptible artefacts. Instead, application of Weber's perceptual mask leads to the invisibility of the watermark. 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. Derivation of a truly optimal detector would involve a likelihood ratio test. For this to be a tractable problem, one has to assume what the probability distribution of s is. For this reason, De Rosa et.al [2] did restrict to a fixed value of s, in their derivation of an optimal detector for multiplicative watermarks. We do not want to make any assumption on the distribution of s, as in practice the value of s is not known, a priori. It is chosen as a trade-off between visibility and robustness of the watermark. Moreover, many attacks against the watermark can be modelled as a transformation of the image affecting the effective embedding depth s. Therefore we have to take a completely different approach, based on maximum likelihood estimation of s.

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), 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.

MAIN RESULT

In the present section we formulate and prove our main result: the derivation of a optimal-estimation based detector for multiplicative watermarks. This is done under the assumption that the original data is modelled by a Rayleigh distribution. We show that under these assumptions, our detector consists of squaring the data and subsequently correlating with the watermark. We do not restrict a priori what values s can take.

The Rayleigh distribution is a suitable model for Fourier coefficients [2]. The Rayleigh probability density function is given by

$$f(x) \;=\; \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}.$$

DCT coefficients are usually modelled by a Weibull probability density function:

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

The Rayleigh distribution is a special case of the Weibull distribution, corresponding to $\beta = 2$ and $\alpha = \sqrt{2}\sigma$. In the third section, we comment on extensions to Gaussian and Weibull distributions.

In the sequel, we use the following notation for inner product and norm of N-tuples:

$$\langle a, b\rangle \;=\; \frac{1}{N}\sum_{i=1}^{N} a_i b_i, \qquad \|a\| \;=\; \sqrt{\langle a, a\rangle} \;=\; \sqrt{\frac{1}{N}\sum_{1=i}^{N} a_i^2}.$$

Theorem 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 Rayleigh distribution with parameter* σ. *Moreover assume that* W *is a known, binary valued sequence (i.e.,* $\forall i : w_i = \pm 1$*), satisfying*

$$\frac{1}{N} \sum_{i=1}^{N} w_i \ = \ 0. \tag{2}$$

The decision variable d *for a detector based on maximum-likelihood estimation of* s *is*

$$d \ = \ \frac{1}{N} \sum_{i=1}^{N} q_i^2 w_i. \tag{3}$$

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 to $d = \frac{1}{N} \sum_{i=1}^{N} q_i^2 w_i$.

Proof: As p_i is distributed according to a Rayleigh distribution with parameter σ, the joint probability density function of $\{q_i\}$, where $q_i = p_i(1 + s w_i)$ and s is viewed as a parameter, is given by

$$f(Q; s) \ = \ \prod_{i=1}^{N} \frac{q_i}{\sigma^2 (1 + s w_i)^2} e^{-\frac{q_i^2}{2\sigma^2 (1 + s w_i)^2}}.$$

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:

$$L(Q; s) \ = \ -2N \log(\sigma) + \sum_{i=1}^{N} log(q_i) - 2 \sum_{i=1}^{N} \log(1 + s w_i) - \frac{q_i^2}{2\sigma^2 (1 + s w_i)^2}.$$

Now, we need to solve

$$\frac{\partial L}{\partial s} \;=\; -\sum_{i=1}^{N} \frac{2w_i}{1+sw_i} + \frac{q_i^2 w_i}{\sigma^2(1+sw_i)^3} \;=\; 0.$$

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

$$\frac{\partial L}{\partial s} \;\approx\; -2\sum_{i=1}^{N} w_i(1-sw_i) + \frac{1}{\sigma^2} q_i^2 w_i(1-3sw_i) \;=\; 0.$$

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

$$\hat{s} \;=\; \frac{\frac{1}{N}\sum_{i=1}^{N} \frac{q_i^2 w_i}{\sigma^2}}{\frac{1}{N}\sum_{i=1}^{N} 3\frac{q_i^2 w_i^2}{\sigma^2} - 2w_i^2}. \qquad (4)$$

Note that $w_i = \pm 1$ implies that

$$\frac{1}{N}\sum_{i=1}^{N} q_i^2 w_i^2 \;=\; \frac{1}{N}\sum_{i=1}^{N} q_i^2 \;=\; \|Q^2\|.$$

This shows that under our conditions

$$\hat{s} \;=\; \frac{\frac{1}{N}\sum_{i=1}^{N} q_i^2 w_i}{3\|Q^2\| - 2\sigma^2}. \qquad (5)$$

■

The requirement in Theorem 1 imposed on the watermark 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, in the same way as for additive watermarks.

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 *Under the assumptions of Theorem 1, the decision variable d satisfies*

$$E[d|s,W] \;=\; 4s\sigma^2,$$

$$\mathrm{var}(d|s,W) = \frac{4\sigma^4}{N}(1+6s^2+s^4).$$

Proof: This proof is a matter of long but straightforward computations. All summations are over the range $1,\ldots,N$. First,

$$\begin{aligned} E[d|s,W] &= E[\frac{1}{N}\sum q_i^2 w_i \; |s,W] \\ &= \frac{1}{N}\sum (1+sw_i)^2 w_i E p_i^2 \\ &= \frac{1}{N}\sum (w_i + 2sw_i^2 + w_i^3) E p_i^2 \\ &= \frac{2s}{N}\sum w_i^2 E p_i^2 \\ &= 4s\sigma^2 \end{aligned}$$

Secondly,

$$\begin{aligned} E[d^2|s,W] &= \frac{1}{N^2}\sum\sum (1+sw_i)^2(1+sw)j)^2 w_i w_j E p_i^2 p_j^2 \\ &= \frac{1}{N^2}\sum\sum (w_i+2sw_i^2+s^2w_i^3)(w_j+2sw_j^2+s^2w_j^3) E p_i^2 p_j^2. \end{aligned}$$

Now denoting $\gamma_0 = Ep_i^2 p_j^2 = 4\sigma^4$ for $i \neq j$ and $\gamma_1 = Ep_i^4 = 8\sigma^4$, we obtain

$$\begin{aligned} E[d^2|s,W] &= \frac{\gamma_0}{N^2}\sum\sum (w_i+2sw_i^2+s^2w_i^3)(w_j+2sw_j^2+s^2w_j^3) \\ &\quad + \frac{\gamma_1-\gamma_0}{N^2}\sum (w_i+2sw_i^2+s^2w_i^3)^2 \\ &= \frac{\gamma_0}{N^2}\left(\sum 2sw_i^2\right)^2 + \frac{\gamma_1-\gamma_0}{N^2}\sum w_i^2 + 6s^2w_i^4 + s^4 w_i^6 \\ &= 4s^2\gamma_0 + \frac{(\gamma_1-\gamma_0)}{N}(1+6s^2+s^4) \\ &= 16s^2\sigma^4 + \frac{1}{N}4\sigma^4(1+6s^2+s^4), \end{aligned}$$

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

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

■

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 quadratic correlation of Theorem 1, we have

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

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

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

and so the quotient would be

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

which differs from the value for multiplicative watermarks by a factor $\frac{\sigma\sqrt{4-\pi}}{\sqrt{2}}$.

GENERALISATION TO OTHER DISTRIBUTIONS

The analysis in this paper was carried out for data p_i which are an i.i.d. sequence drawn from a Rayleigh distribution. The result can easily be generalized to data drawn from a Gaussian distribution. In this case, we obtain

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

The proofs remain virtually unchanged.

With a little bit more work, the result can also be extended to Weibull distributions. If we assume that the data p_i are i.i.d. distributed according to the Weibull distribution, then we can carry out the same derivation of a detector based on maximum-likelihood estimation of s. This leads to the decision variable

$$d \;=\; \frac{1}{N}\sum_{i=1}^{N} q_i^{\beta} w_i.$$

Note that the Rayleigh distribution with parameter σ is a special case of the Weibull distribution, corresponding to the parameters $\beta = 2$, $\alpha = \sqrt{2}\sigma$.

CONCLUSIONS

In this paper we have derived an optimal-estimation based detector for multiplicative watermarks. Our derivation shows that the sample values should be squared before correlation. The resulting detector differs from those used nowadays, which are based on correlating the samples directly with the watermark. This result shows that very likely it is possible to improve on the common practice to use correlation of a watermark with (non-squared) data to detect additive watermarks which are embedded using perceptual masks (the result of which embedding can sometimes be modelled by a multiplicative watermark).

Our result has as a consequences for implementation, that before anything is done all sample values have to be squared. If, for instance, the watermark detection uses tiling, this has to be done after squaring.

REFERENCES

[1] I.J. Cox, J. Killian, F. Thomson Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Trans. IP*, 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 *Proc. 3rd Int. Inform. 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. *Proc. 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 *Proc. 3rd Int. Inform. Hiding Workshop*, pages 303–314, Dresden, 1999.

[5] M.L. Miller and J.A. Bloom. Computing the probability of false watermark detection. In *Proc. 3rd Int. Inform. 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 *Proc. 3rd Int. Inform. Hiding Workshop*, pages 219–244, Dresden, 1999.