WATERFILLING: A CONNECTION BETWEEN INFORMATION THEORETIC AND PERCEPTUALLY BASED WATERMARKING

Siva Somasundaram and R. Chandramouli

Multimedia Systems, Networking, and Communications (MSyNC) Lab
Department of Electrical and Computer Engineering
Stevens Institute of Technology

ABSTRACT

We present a hybrid approach to image watermarking that exploits results from both information theory and perceptual studies. Towards this purpose we use a waterfilling-type algorithm to establish a connection between information theoretic and perceptually based watermarking. Information theoretic capacity formula for a Gaussian signal over a Gaussian channel is used and some parameters of the corresponding waterfilling-type algorithm are chosen based on a perceptual criterion. Image sub-blocks to be watermarked are chosen based on local statistics of each sub-block and the embedded watermark strength for each of these sub-blocks are computed based on a perceptual criterion. Towards this end an iterative algorithm is proposed. Experimental results are presented to discuss the trade-off between information theoretic capacity, robustness and perceptual distortion.

1. INTRODUCTION

Information theory has been a valuable tool in studying watermarking problems. Its applications include a definition of watermarking capacity [1,2], capacity computation for watermarking channels [1,2,3], definition of watermarking security and many more. Many rich set of mathematical models and tools form the basis of information theoretic studies of watermarks. A parallel approach to watermarking is perceptually based [4,5,9]. Perceptual models take into consideration the human audio-visual system while information theoretic approaches do not attempt this. Therefore there has been a debate about the pros and cons of information theoretic versus perceptually based approaches to studying watermarking. In [10], a watermarking technique that uses both information theoretic and perceptually criteria has been proposed. We extend these results in this paper.

Based on statistical models for the watermark signal, host signal and attacks (if any), information theory predicts the maximum amount of watermark information that can be embedded in a host signal. Therefore, a wide class of host signals that have similar statistical characteristics will be treated the same way in this method. While this is helpful in computing the mathematical capacity for this class of signals, it must be noted that the individual signals in this class may themselves possess significantly different perceptual characteristics. Since, for many watermarking applications perceptual distortions must be kept a minimum, the information theoretic capacity may not be accurate from a perceptual sense. Therefore, it is argued by practitioners that information theoretic methods may not be very helpful from a perceptually based real-life application perspective, especially in terms of watermarking capacity. We note with caution that perceptual metrics are themselves controversial and not unique. Moreover, to compute the watermarking capacity from a perceptual point of view, each host signal has to be treated separately thus making this process tedious. It is our opinion that (arguably) perceptual metrics lack the generalization capability of information theory.

This paper is an attempt to extend some of the results presented in [10]. In [10], a perceptually based waterfilling algorithm for watermarking is proposed. However, the local statistics of a watermarked image sub-block are not exploited. Instead, a global average statistic is used. This leads to situations where textured areas of an image are not watermarked, etc. Therefore, in this paper we improve upon this result by choosing image sub-blocks for watermarking based on local statistics computed using its 8-neighborhood sub-blocks. Note that, traditionally, waterfilling deals with optimally allocating energy for various sub-channels in a communication system that operates under a total energy constraint [6] such that capacity is achieved. Independence of the sub-channels is a typical assumption during this process. We extend this idea for image watermarking by computing the watermark energy allocation based on a perceptual criterion for the various sub-channels (image sub-blocks) derived from an image such that the information theoretic watermarking capacity is achieved.

The paper is organized as follows. Section 2 describes the watermarking capacity computation and the proposed waterfilling-type algorithm that achieves capacity while maintaining the perceptual distortion constraint. Experimental results are given in Section 3 followed by concluding remarks in Section 4.

2. INFORMATION THEORETIC CAPACITY AND PERCEPTUALLY BASED WATERFILLING

In this section we first describe the watermark embedding technique and watermarking capacity for a parallel watermark channel model followed by the proposed waterfilling algorithm.

2.1 Watermark Embedding and Capacity

We employ a variation of the spread-spectrum watermark [8] in the spatial domain and assume blind watermark detection where the detector has access to the watermarking key and other relevant side information. Therefore, the watermark channel uncertainty consists of the host image and other attacks. For the current treatment, we ignore uncertainties due to attacks since our primary aim is to introduce a technique that utilizes information theoretic and perceptual analysis in watermark embedding; however, we note that any attack(s) on the watermark can be incorporated into the proposed analysis via appropriate mathematical models without much additional effort.

Before embedding the host image is first partitioned into non-overlapping sub-blocks and without loss of generality let $X_t = \{x_1, x_2, ..., x_L\}$ denote the local mean subtracted pixel values in such a sub-block, $t = 1, 2, ...N$. Further, let $x_i \sim \mathbf{N}(0, \sigma_k^2)$ be Gaussian distributed. If sub-block k is chosen for embedding, then,

$$y_i = x_i + w_{i,} \quad i = 1,2,...,L \tag{2.1}$$

where $w_i \sim \mathbf{N}(0, \gamma^2_k)$ is a zero-mean, Gaussian distributed watermark with variance γ^2_k. Note that each watermarked sub-block here can have a different watermark energy or variance. The embedding in Eq. (2.1) is implemented using the following equivalent form,

$$y_i = x_i + \alpha_k \tilde{w}_i \tag{2.2}$$

where $\tilde{w}_i \sim \mathbf{N}(0,1)$ and α_k is a scale parameter that determines the watermark energy for that sub-block, k. If M out of the N sub-blocks are selected for watermark embedding, then treating this as M independent parallel channels, it can be shown that [6] the watermarking capacity is given by

$$C = \frac{1}{2} \sum_{k=1}^{M} \log_2 \left(1 + \frac{\gamma^2_k}{\sigma^2_k} \right) \tag{2.3}$$

where $\sum_{k=1}^{M} \gamma^2_k \le E$ is the total watermark energy constraint. But, we note that unlike a communication system where the energy constraint E can be specified *a priori* based on physical system limitations, in watermarking the value for E is not obvious. A good estimate for E depends on a variety of factors such as robustness, perceptual distortion and the total capacity requirement. Therefore, we concentrate on choosing the values for γ_k's and compute the resultant value of E. This is a deviation from the traditional information theoretic waterfilling. Also, note from Eq. (2.3) that the image sub-blocks with smaller values of σ^2_k possess higher capacity. We exploit this fact using a perceptual metric. The issues we address are the following:

- Which image blocks to choose for watermark embedding from a perceptual perspective such that capacity is achieved?
- What must be the watermark energy allocated to the chosen blocks based on a perceptual constraint?

We propose a solution in the next subsection based on a waterfilling-type idea rooted in information theoretic and perceptual arguments.

2.2 Waterfilling-type Algorithm for Watermarking

Perceptually based watermarking [5] predicts that features of an image that are perceptually significant must be chosen for watermark embedding. The just noticeable difference criterion is usually used for this purpose. This criterion is typically computed in the transform domain. The discrete cosine transform based spread spectrum watermarking [8] technique embeds watermark in the mid-frequency components that are perceptually significant. In [5] transform coefficients that pass a threshold test based a perceptual model are selected for watermark insertion. In most perceptually based watermarking methods, the notion of capacity achieving embedding has not been considered. Reasons for this include the following:

- Unlike information theoretic capacity, there seems to be no closed form solution to the *perceptual watermarking capacity*.
- A good perceptual metric is itself an active area of research and different perceptual metrics could give rise to different notions of perceptual capacity.

Therefore, the natural question to ask is: can we use the information theoretic capacity measure in conjunction with a perceptual metric? We show that this is possible and the solution is consistent with both information theory and perceptual studies. Before embarking towards this goal, we make some useful observations. If E is large then the capacity is also large and the watermark is robust against attacks at the cost of increased perceptual distortion. On the other hand, a small value of E results in a smaller perceptual distortion, smaller capacity and a not so robust watermark. Note that a choice of E is also image dependent. Therefore, we indirectly compute E as previously stated. Now, we state an information theoretic result for achieving capacity [6]:

$$\gamma^2{}_t = \begin{cases} T_t - \sigma_t^2, \text{ if } \sigma_t^2 < T_t \\ 0, \text{ if } \sigma_t^2 \ge T_t, \ t = 1,2...N \end{cases} \tag{2.4}$$

Note that N-M of the γ_t's will be equal to zero if only M image sub-blocks are chosen for watermark embedding. Clearly, the choice of T_t controls which sub-blocks are chosen for embedding. We choose T_t as the median of the variances of the 8-neighbouring blocks of a candidate image sub-block. Note that this measure is in the same spirit as some of the other perceptually based metrics [4,5,10] used for watermarking. This choice of T_t seems to be also consistent with the observation made earlier that an image sub-block with a lower local variance has a higher capacity. Our experimental results also seem to corroborate this observation. Therefore, if the variance of a candidate image sub-block is greater than or equal to T_t then that block is not chosen for embedding as it will violate the perceptual distortion constraint. Once the image sub-blocks for embedding are chosen the next step is to compute the watermark energy allocated to each of these blocks consistent with perceptual quality metrics. This process is developed as an iterative procedure. The steps involved in the proposed hybrid (information theory+perceptual criterion) watermarking approach are summarized below:

Hybrid Watermarking Algorithm:

```
Divided an image into 8x8 candidate sub-blocks.
Choose the first (candidate) block:
Step 1:
    $T_t$ = median{variances of 8-neighbouring sub-blocks}
    If variance of candidate block is greater or equal to $T_t$
    choose next candidate block and repeat Step 1.
    If not, move to Step 2.

Step 2:
    Embed the Gaussian distributed watermark with
    initial total scale factor $\alpha_k^{(1)} = (\sqrt{T_t - \sigma_t^2})$,
    i.e., $\alpha_i^1 = \alpha_k^1 / 64$ for all pixels ($i = 1,2,...,64$)
    in that block.

Step 3:
    Compute the current perceptual distortion metric ($\mathfrak{C}^1$)
    in the LL band for the watermarked sub-block in the
    wavelet domain as given in [4]. Let $\mathfrak{M}$ be the
    maximum allowable distortion in the LL band [4].

Step 4:
    While ( $\mathfrak{C}^n > \mathfrak{M}$ )
        $\alpha_k^{(n+1)} = \alpha_k^{(n)} + \theta\left(\frac{\mathfrak{C}^{(n)} - \mathfrak{M}}{\mathfrak{M}}\right),\ n = 1,2,...$
        where $0 < \theta < 1$. Embed watermark in current
        candidate block with the value $\alpha_k^*$ after convergence.
    Choose the next block for embedding and go to Step 2.
```

It is easy to prove that Step 4 of the algorithm converges.

3. EXPERIMENTAL ANALYSIS

For the experimental analysis of the proposed algorithm, 25 images were chosen carefully so that they exhibit a wide variety of image characteristics. Image samples included those with dark backgrounds, light backgrounds, highly textured features, contrasting letters over the background, and paintings. The proposed watermark energy computation technique was tested on these 25 images. We only present the results obtained for two images *Fish* and *Lena* here due to space limitations. The image sub-block size was chosen to be equal to 8x8. Note that the edge blocks will not be chosen for embedding by our algorithm since they do not have 8-neighbour sub-blocks.

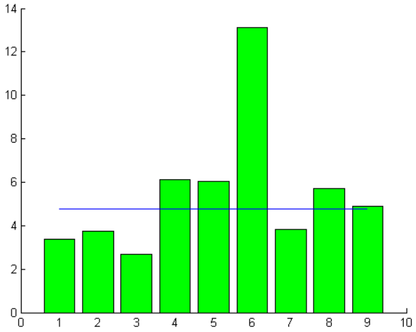

Figure 1: Variances of 8-neighborhood sub-blocks and the candidate sub-block (no. 9 on X-axis). Median is shown as the horizontal line. This candidate block is not chosen for watermark embedding.

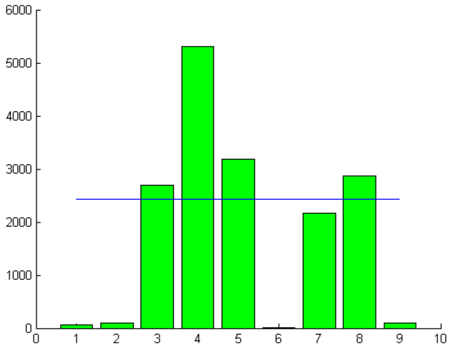

Figure 2: Variances of 8-neighborhood sub-blocks and the candidate sub-block (no. 9 on X-axis). Median is shown as the horizontal line. This candidate block is chosen for watermark embedding.

Fig. 1 and 2 each depict an example of the local statistic of a candidate block in Lena. In Fig. 1 the candidate block is not chosen for embedding because it fails the threshold test and in Fig. 2 the candidate block passes the threshold test and hence is chosen for watermark embedding. Figs. 3 and 4 show the watermarked images. Clearly, the watermark has not perceptually degraded image. Also shown in these images are the white squares corresponding to the blocks in the image that has been watermarked. Note that the sub-blocks to be watermarked are chosen based on the variance map. For example, in the Fish image the sub-blocks corresponding to the transition regions where changes in the variance from one region to the next is significant are not chosen. This is also consistent with similar observations made in [11]. Fig. 5 shows the convergence of the watermark energy estimation algorithm (Step 4 in the watermarking algorithm) for sub-block number 2934 when

$\mathfrak{M} = 14.4$. Speed versus accuracy is traded-off by the choice of the θ parameter. A lower value produced an accurate estimate of the watermark scale factor at the cost of an increased number of iterations.

Figure 3: Watermarked blocks for the Lenna image are shown as white squares. These blocks and the watermark energy were chosen based on the proposed algorithm.

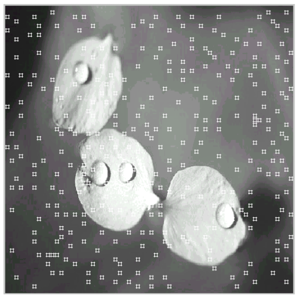

Figure 4: Watermarked blocks for the Fish image are shown as white squares. These blocks and the watermark energy were chosen based on the proposed algorithm.

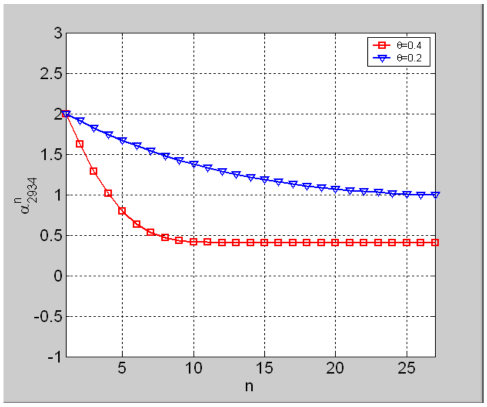

Figure 5: Convergence of watermark energy estimation algorithm.

4. CONCLUSION

A hybrid watermarking algorithm based on information theoretic and perceptual metrics is proposed. Information theoretic capacity maximization is the goal of this algorithm while preserving perceptual quality. A method to identify image sub-blocks for watermark embedding based on local statistic is proposed. This is then followed by the introduction of an iterative algorithm for watermark energy allocation that satisfies a perceptual criterion. Experimental results show that the proposed method achieves its objectives. The results are also seen to be consistent with observations made by other research efforts with goals consistent with ours.

5. REFERENCES

[1] R. Chandramouli, "Data hiding capacity in the presence of an imperfectly known channel," SPIE *Proceedings of Security and Watermarking of Multimeida Contents II*, Vol.4314, 2001.

[2] P. Moulin and M.K. Mihcak ,``The data hiding capacity of image sources," *preprint,* available at http://www.ifp.uiuc.edu/~moulin/paper.html

[3] J.J.K. O' Ruanaidh, W.J. Dowling, and F.M. Boland, ``Watermaking digital images for copyright protection." *IEE Proc. Vis. Image Signal Process* , vol 143, no.4., August 1996.

[4] L.C. Yung and S.F. Chang , "Watermakring capacity of digital images based on domain specific masking effects," *Proc. International Conference on Information Technology*, Las Vegas, pp. 90-94, 2001.

[5] R. Wolfgang, C. Podilchuk, and E.J.Delp, "Perceptual watermarks for digital images and video," *Proceedings of IEEE*, vol.87, no.7, July 1999.

[6] Cover and Thomas "Information theory and coding" , John Wiely & sons, 1991.

[7] A.B. Watson, "DCT quantization matrices visually optimized for individual images", *Proc. SPIE*, vol.1913, pp. 202-216, 1993.

[8] I.J. Cox, J. Kilian, T. Leighton and T. Shamoon, ``Secure spread spectrum watermarking for multimedia'', *IEEE Trans. on Image Processing, 6, 12, 1673-1687,* (1997).

[9] Barni, M.; Bartolini, F Piva, ``Improved wavelet-based watermarking through pixel-wise masking," *IEEE Ttrans. on Image processing* , pp. 783 –791, vol 10., no.5, 2001.

[10] S. Somasundaram and R. Chandramouli, ``Perceptually based waterfilling for watermarking," Special session on Multimedia Watermarking, Proc. ISCAS 2002, Arizona (to appear).

[11] I. Cox, M. Miller and J. Bloom, *Digital Watermarking,* Morgan Kauffman, 2001.