DESIGNING ROBUST WATERMARKS USING POLYNOMIAL PHASE EXPONENTIALS

Bijan G. Mobasseri, Yimin Zhang, Moeness G. Amin, Behzad M. Dogahe

Center for Advanced Communications
Villanova University
Villanova, PA 19085

ABSTRACT

In this paper, we propose a known-host-state methodology for designing image watermarks that are particularly robust to compression. The proposed approach outperforms traditional spread spectrum watermarking across all JPEG quality factors. The fundamental approach uses 2D chirps as spreading functions, followed by chirp transform, to recover the watermark. Because this method can spectrally shape the chirp to match image content and JPEG quantization, its performance is greatly enhanced. The energy localization of the chirp is exploited to embed low power watermark per image blocks while maintaining reliable detection performance.

1. INTRODUCTION

Digital watermarking is the process of securely embedding invisible signatures within a cover media with no perceived impact. Developing a new watermarking algorithm requires the definition of five components: 1): cover media, 2): watermark, 3): embedding and extraction, 4): perceptual metric, and 5): resilience and security criteria. Watermarking has been implemented in spectral as well as spatial domains. Arguably, the best known watermarking technique is spread spectrum (SS). Cox made one of the earliest references to SS watermarking [1]. Hernandez et al. [2]applied the same idea to 8x8 block DCT transformation of images, closely following the JPEG standard. Their approach is similar to the spatial domain SS watermarking, which does not resort to masking models, proposed earlier by Hartung and Girod [3].

The concept of spread spectrum can be applied to spectral as well as spatial domains. In [4], each watermark bit is spread by a 2D modulation function and added to nonoverlapping sets of image pixels driven by a density metric. This is similar to the phase dispersion method proposed in [5]. The two approaches in [4,5] use the same model to spread watermark bits. While one uses a PN sequence, the other designs a carrier with a flat spectrum and pseudorandom phase. SS watermarking is then an attempt to find the proper spreading function.

In this paper, we use polynomial phase exponentials, specifically, a chirp function, as the spreading function. Chirps bring three properties to the table. First, chirp signals allow for tuning and spectral shaping of the watermark in a way that traditional spread spectrum watermarking using PN sequences does not. Second, as a highly localized signal, chirp/watermark energy can be spread out in the image and then integrated at detection. This allows for low power watermark embedding on a local basis. Third, there has been considerable work in time-frequency processing techniques in the areas of speech, communications, fault structures, automation, biomedicine, radar, and sonar. These techniques provide easily accessible information about the signal spectral localization over short time periods and spatial segments [6]. We apply the chirp transform and matched filter processing in order to design and detect the chirp characteristics suitable for watermarking. The chirp transform applied in this paper does not use the fast computations offered by the discrete chirp-Fourier transform in [7]. In a prior work, Stankovic et al. [8] used chirps as digital watermarks by adding a single chirp to the entire image. This algorithm is best suited to copyright and ownership verification applications. The ability to embed and detect different chirps per image block, however, allows for data hiding applications where the extracted watermark may be an information-bearing bitstream. Another point of departure from [8] is the exploitation of *known-host-state-methods* [9]. This approach was first suggested by Cox as a communication problem with side-information [10], and was based on Costa's dirty paper writing [11]. We incorporate this idea into our work by tuning the chirp. This observation is in marked contrast to spread spectrum watermarking, where the spreading function, in the form of a PN sequence, is unrelated to host signal statistics.

2. RATIONALE FOR A NEW WATERMARK

We follow the watermarking model in [4]. We embed p bits $B=\{b_o,b_1,...,b_{p-1}\}$ in image $I(x,y)$. For each bit b_i

This research was supported in part by ONR grant N00014-04-1-06-30.

we define a 2D modulation function defined over set of pixels S_i

$$s_i(x,y)=\begin{cases}p_i(x,y) \; if\,(x,y)\in S_i \\ 0\end{cases} \tag{1}$$

$p_i(x,y): PN\ sequence$

where $S_i \bigcap S_j = \emptyset, i \neq j$. The watermarked image is now defined by

$$I_w(x,y)=I(x,y)+w(x,y) \tag{2}$$

where

$$w(x,y)=\sum_{i=0}^{p-1}(-1)^{b_i}s_i(x,y)\alpha(x,y) \tag{3}$$

$\alpha(x,y)$ controls watermark strength. The above model spreads each watermark bit by a PN sequence and then additively modifies image pixels over the defined region of the image. Watermarks are detected by a classical correlation detector, provided the decoder has access to the seed of the PN sequence, presumably communicated via a secret key exchange protocol. The proposed model, although effective, is suboptimal in the sense that no host-state-statistic is taken into account. In this work, we propose a class of watermarks that are *host-aware*.

Robustness to compression is a basic requirement in watermarking. In order to achieve such robustness, we suggest that the watermark be spectrally shaped to escape JPEG. Since the high frequency suppression of the JPEG standard is well-known, it should be possible to design a watermark that will be unaffected by compression.

3. WATERMARK EMBEDDING AND DETECTION

Partition an *NxN* image into *M* square blocks. A complex 2D chirp is defined as follows:

$$W(x,y)=e^{j\pi\left(\beta_x x^2+\beta_y y^2\right)+j2\pi\left(f_x x+f_y y\right)}\{(x,y)\in 0,M-1\} \tag{4}$$

where β_x and β_y are chirp rates, and f_x and f_y are initial spatial frequencies. For the rest of this paper, we use a single pair $\{\beta, f\}$. Spectral shaping can now take place by adjusting the pair $\{\beta,f\}$ (Fig. 1). Following the watermarking model in (3), the image block located at pixels (m,n) is watermarked as follows:

$$I_w(m,n,x,y)=I(m,n,x,y)+k\,\mathrm{Re}\big[d(m,n)W(x,y)\big] \tag{5}$$

where [] stands for the integer part, $d(m,n)$ is the watermark bit drawn from *B*, and *k* controls the PSNR in the watermarked image. Watermark detection is based on 2D chirp transform defined in (6). To recover *B*, the decoder requires knowledge of the specific pair $\{\beta_o,f_o\}$ of the embedded chirp. This pair can be obtained in the chirp transform domain by seeking the peak of $C(m,n,\beta,f)$. This peak can be enhanced by integrating $C(m,n,\beta,f)$ over all image blocks followed by peak searching in (7).

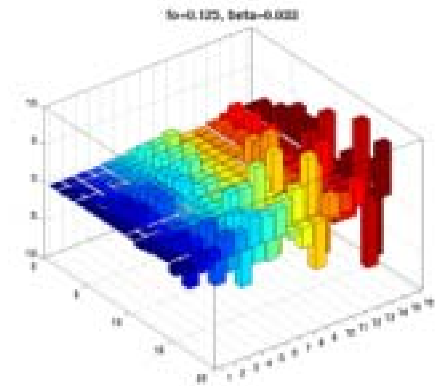

Figure 1- DCT of a 16x16 chirp.

$$\begin{aligned}C(m,n,\beta,f)&=\sum_{x=0}^{M-1}\sum_{y=0}^{M-1}I_w(m,n,x,y)U^*(x,y,\beta,f)\\&=\sum_{x=0}^{M-1}\sum_{y=0}^{M-1}I(m,n,x,y)U^*(x,y,\beta,f)\\&+\sum_{x=0}^{M-1}\sum_{y=0}^{M-1}k\,\mathrm{Re}\big[d(m,n)W(x,y)\big]U^*(x,y,\beta,f)\end{aligned} \tag{6}$$

with

$$U(x,y,\beta,f)=e^{j\pi\beta\left(x^2+y^2\right)+j2\pi f(x+y)}$$

$$C(\beta,f)=\sum_m\sum_n\left|C(m,n,\beta,f)\right| \tag{7}$$

Eq(7) allows us to distribute watermark power over the entire image and then integrate. This power distribution makes watermark detection by unauthorized users more difficult because each block alone carries an amount of power insufficient for reliable detection. Note that $C(\beta,f)$ for an unmarked image peaks at $\{\beta,f\}\cong 0$. To prevent others from performing identical peak detection and recovering the same information, the following procedure is implemented. Instead of embedding the same chirp in every block, we draw from a family of $\{\beta,f\}$ and embed different pairs in different blocks. This association is then communicated to the decoder via secure key exchange. Unless this key is known and image blocks are de-chirped with correct $\{\beta,f\}$, the attacker will not observe a peak in (7). The embedded bit can now be recovered from the sign of $C(m,n,\beta_o,f_o)$ as follows:

$$\hat{d}(m,n)=\begin{cases}1 & \mathrm{Re}\{C(m,n,\beta_o,f_o)\}\geq 0\\ 0 & \mathrm{Re}\{C(m,n,\beta_o,f_o)\}<0\end{cases} \tag{8}$$

The bit error rate (BER) is controlled by the relative strength of projection of the chirp on the image vs. the chirp itself. Let

$$I_w = I + kdW_{\beta f}, d \in \{+1,-1\} \quad (9)$$

Decoder output is given by

$$\begin{aligned} &< I_w, W_{\beta f} > = < I, U_{\beta f} > + kdE_{\beta f} \\ &where \\ &E_{\beta f} = < W_{\beta f}, W_{\beta f} > \end{aligned} \quad (10)$$

Correct detection is guaranteed if $\left| < I, U_{\beta f} > \right| < E_{\beta f}$. Probabilities of error are given by (11). Not every block needs to be watermarked. For example, a block that is slated to carry +1 and satisfies $< I, U_{\beta f} >> kE_{\beta f}$ is left alone since the decoder will decide in favor of +1 regardless.

$$\begin{aligned} P(e|s=-1) &= P(< I, U_{\beta f} > \ > kE_{\beta f}) \\ P(e|s=+1) &= P(< I, U_{\beta f} > \ < -kE_{\beta f}) \end{aligned} \quad (11)$$

4. JPEG COMPRESSION AND SPECTRAL SHAPING

Below, we show the flexibility of using chirp over traditional SS watermarking in compressive environments. SS watermarking offers substantial robustness to compression [12]. This robustness is achieved through the available processing gain. Increasing processing gain in spread spectrum watermarking, however, reduces the embedding rate. This reduction occurs because higher processing gains can only be achieved by using larger image blocks. In chirp-based watermarking, robustness to compression is achieved in an entirely different manner. To prevent the JPEG from removing the watermark, we spectrally shape the chirp to make it survive compression. This shaping can be achieved by varying $\{\beta, f\}$ and monitoring BER. In contrast, the PN sequence in SS watermarking has no such tuning capability and retains a white spectrum regardless.

The key issue is the chirp selection, which survives a specific compression factor. Rewrite (5) as follows:

$$I_w = I + kdW_{\beta f}, d \in \{+1,-1\} \quad (12)$$

Define JPEG quantization matrix by $Q = [q_{ij}], \{i, j = 1,...,8\}$. Quantized DCT coefficients of watermarked image block are given by

$$\left\lceil \frac{dct(I_w)}{Q} \right\rceil = \left\lceil \frac{dct(I)}{Q} + \frac{dct(ksW_{\beta f})}{Q} \right\rceil \quad (13)$$

where $\lceil \ \rceil$ designates rounding to the nearest integer. Division in (13) is a term-by-term division of two 8x8 matrices. The dequantized image block is given by

$$I_w^* = dct^{-1}\left(Q \times \left\lceil \frac{dct(I)}{Q} + \frac{dct(ksW_{\beta f})}{Q} \right\rceil \right) \quad (14)$$

I_w^* is then used in (6). The watermark may be eliminated by compression through quantization. Since quantization is a nonlinear operation and is the sum of image and quantized watermark components, the DCT values of the chirp alone, without considering the content of the image, do not determine watermark survival. A watermark can be considered removed if DCT coefficients are quantized to the same value, with or without the watermark. Since both image and watermark are available to the encoder, it is possible to ensure watermark survival by choosing appropriate $\{\beta, f\}$ pairs. A finer point here is that watermark survival is not absolute; there are different degrees of watermark content in I_w^*. This is because there are actually 64 terms in (13). Theoretically, even if one frequency out of 64 retains the watermark, the watermark has survived but may not be reliably detected, resulting in a large BER.

We can quantify the degree of watermark survival using the following measure:

$$e = \frac{1}{M} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \left| I_w^*(i,j) - I^*(i,j) \right| \quad (15)$$

$I^*(i,j)$ is an unmarked compressed image block. If this difference is zero, the watermark is entirely removed by compression. For a fixed PSNR and compression ratio, e is a function of $\{\beta, f\}$. Chirp design amounts to selecting the pair that results in large e.

5. EXPERIMENTS

Our test image was *Lena* in grayscale. The 512x512 image was divided into 16x16 blocks for a total 1024 blocks. The embedding capacity for this image was 1024 bits. In order to tune the chirp to the image and a range of JPEG quality factors, we computed BER contours at the encoder. Figure 2 shows BER contours for various $\{\beta, f\}$. The horizontal and vertical axes were $\{\beta, f\}$, respectively.

BER contours can be used to pick $\{\beta, f\}$ pairs that meet specific BER requirements. Note that many different chirps can be used to achieve the same BER requirement. Thus, for security purposes, we can use different chirps that still provide the same BER. Figure 3 shows two watermarked images, both with acceptable quality. Figure 4 shows BER performance for chirp and spread spectrum watermarking respectively. When selected as the spreading function, chirp outperforms m-sequences across all Q-factors. We see similar trends for 5 other test images we have run.

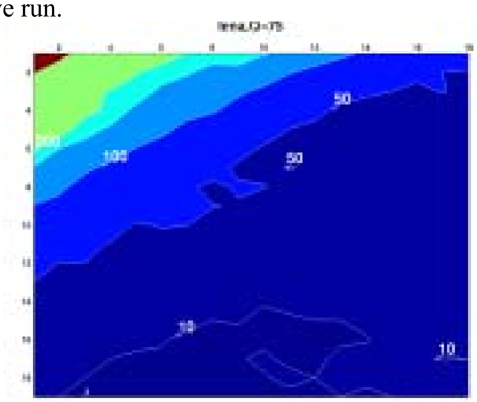

Figure 2- *BER contours for JPEG quality factors 75. The numbers indicate bit errors out of 1024 embedded bits.*

Figure 3- *Two watermarked images using chirp(left) and spread spectrum(right). Both carry 1024 bits at PSNR=40dB.*

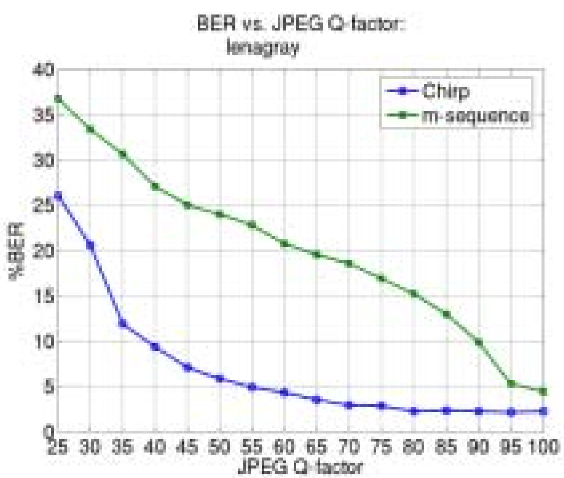

Figure 4-BER vs JPEG quality factor. Chirp consistently outperforms m-sequences by a wide margin.

6. CONCLUSIONS

We used 2D chirp functions for image watermarking. The flexibility provided by the chirp allowed us to tune the chirp in unprecedented ways. Performance advantages over spread spectrum technique were demonstrated.

7. REFERENCES

[1] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” *IEEE Trans. Image Processing*, vol. 6, pp. 1673–1687, Dec. 1997.

[2] J.R. Hernandez, M. Amando and F. Perez-Gonzalez, “DCT-domain watermarking techniques for still images: detector performance analysis and a new structure,” *IEEE Trans. Image Processing*, vol. 9, pp. 55-68, Jan. 2000.

[3] F. Hartung and B. Girod, “Digital watermarking of uncompressed and compressed video,” *Signal Process.*, vol. 66, pp. 283–301, May 1998.

[4] M. Kutter and S. Winkler, “A vision-based masking model for spread-spectrum image watermarking, “*IEEE Trans. Image Processing,* vol. 11, pp. 16-25, Jan. 2002.

[5] C. Honsinger and M. Rabbani, “ Data embedding using phase dispersion,” Eastman Kodak, 2000.

[6] A. Papandreou-Suppappola, *Applications in Time-Frequency Signal Processing*, CRC Press, 2003.

[7] X. Xia, “Discrete chirp-Fourier transform and its application to chirp rate estimation,“ *IEEE Transactions on Signal processing,* pp. 3122-3133, November 2000.

[8] S. Stankovic, I. Djurovic and I. Pitas, “Watermarking in the space/spatial-frequency domain using two-dimensional Radon–Wigner distribution,” *IEEE Trans. Image Processing*, vol. 10, pp. 650-658, April 2001.

[9] F. Pérez-González*,* F. Balado, and Juan R. Hernández Martin*, “*Performance analysis of existing and new methods for data hiding with known-host information in additive channels,” *IEEE Trans. Image Processing*, vol. 51, pp. 960–980, April 2003.

[10] I. J. Cox, Matthew L. Miller, and Andrew L. McKellips, “Watermarking as communications with side information,” *Proc. IEEE*, vol. 87, no. 7, pp. 1127–1141, Jul. 1999.

[11] Max H. M. Costa, “Writing on dirty paper,” *IEEE Trans. Information Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[12] B. G. Mobasseri, *“*A spatial video watermark that survives MPEG,*” IEEE International Conference on Information Technology: Coding and Computing, Las Vegas, March 27-29, 2000.*