# Image Authentication and Restoration Using Irregular Sampling for Traffic Enforcement Applications

Xunzhan Zhu, Anthony T. S. Ho
School of Electronics and Physical Science
University of Surrey
Guildford, GU2 7XH, UK
{x.zhu; a.ho}@surrey.ac.uk

Pina Marziliano
School of Electrical and Electronic Engineering
Nanyang Technological University
Nanyang Avenue, 639798, Singapore
epina@ntu.edu.sg

## Abstract

*Traffic image sequences are important information for the purpose of traffic system control and traffic accidents surveillance, for example, the determination of offending vehicles. A semi-fragile watermarking method for the automatic authentication and restoration of traffic images using irregular sampling is described. Watermarks are embedded into the pinned field of the pinned sine transform (PST) of the original image, which reflects local malicious tampering on the texture of the image. When tampered blocks are detected, the restoration problem is formulated as an irregular sampling problem in approximation subspaces. These blocks are then reconstructed, making use of the information embedded in the same watermarked image, through iterative projections onto convex sets in approximation subspaces. The restoration process is a variant of the Papoulis-Gerchberg algorithm, and is robust to common image processing operations such as lossy transcoding and image filtering.*

## 1. Introduction

Image processing techniques applied to automatic road traffic enforcement have been widely investigated by many researchers, [3, 4]. Image sequences captured by vision-based cameras are important tools in the planning, maintenance, and control of any modern transport system, for example, the parameter (volume, speed, type of vehicles and so on) evaluation of the traffic flow, or the determination of offending vehicles during traffic accidents. Therefore, there is a need for verification or authentication of the integrity of the image content. And in some applications, approximate restoration of the tampered portions is also desirable. A semi-fragile watermarking is a potential solution to the image content authentication problem which seeks to verify that the content of the multimedia has not been modified by any of a predefined set of illegitimate distortions, while allowing modification by legitimate distortions.

In this paper, we propose a novel semi-fragile watermarking method for the authentication and self-restoration of traffic images, which is an extension of our previous work in [1]. The watermark signal is generated by combining a pseudo-random signal with some prior knowledge of the carrier image, which are convex sets in some approximation subspace. This mixed signal is then embedded into the "pinned field" of the image. When tampered areas are detected, they are restored as irregular sampled signal in approximation subspaces [2]. Experimental results showed that accurate tamper detection and restoration were still possible after lossy transcoding and other common image processing operations.

In the next section, a short review of irregular sampling in approximation spaces is presented. In Section 3 we introduce our proposed approach. The simulation results are reported in Section 4, followed by concluding remarks in Section 5.

## 2 Restoration from irregular samples

In this paper, we address the problem of detecting tampered blocks in a watermarked image and then restoring them. Assume that we have no prior knowledge of the corruption channel, the problem can be formulated as obtaining an incomplete set of data with lost packets, which can be cast as an irregular sampling problem.

The irregular sampling problem can be solved by projection onto convex sets (POCS) method [2] under the assumption that the signal belongs to two linear convex sets with non-empty intersection. It involves the iterative implementation of projection $\mathcal{P} = \mathcal{P}_b\mathcal{P}_a$, where the first projection $\mathcal{P}_a$ is onto a band-limited subspace $l_a$ and the second projection $\mathcal{P}_b$ is onto the space of unknown samples $l_b$. In [2],

Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC'06)
0-7695-2616-0/06 $20.00 © 2006 IEEE

PG variants were investigated in different linear and nonlinear approximation subspaces and it was showed that the basis which approximates a given signal better gives a better reconstruction.

In this paper, we define the basis vectors $\{\mathbf{g}_m\}$ to be the basis vectors of discrete cosine transform, which is considered the best approximation of natural images. To facilitate the watermarking based restoration problem, we extend the nonlinear approximation method by re-defining the convex sets:

(1) $l_a$ denotes the subset of $\mathcal{H}$, where $\mathcal{H}$ is a Hilbert space, composed of all functions whose cosine transform coefficients satisfy the constraint

$$\mathrm{sgn}(F(u,v)) = \Gamma(u,v) \tag{1}$$

in a prescribed region $\Delta$ of the frequency domain, with

$$\mathrm{sgn}(y) = \begin{cases} 1 & y \geq 0 \\ 0 & y < 0 \end{cases},$$

where $F(u,v)$ is the DCT coefficient of $f(x,y)$, and $\Gamma(u,v) \in \{0,1\}$ is a known binary function.

The projection of an arbitrary $f \in \mathcal{H}$ onto $l_a$ is realized by

$$\mathcal{P}_a f \leftrightarrow \begin{cases} F(u,v) & (u,v) \in \Delta,\ \mathrm{sgn}(F(u,v)) = \Gamma(u,v) \\ 0 & (u,v) \in \Delta,\ \mathrm{sgn}(F(u,v)) \neq \Gamma(u,v) \\ F(u,v) & (u,v) \notin \Delta \end{cases}. \tag{2}$$

(2) $l_b$ denotes the set of all functions in $\mathcal{H}$ which assume prescribed values $\mathbf{\Theta}$ over a closed region $\Delta$. The projection onto $l_b$ is realized by

$$\mathcal{P}_b f = \begin{cases} \Theta(x,y) & (x,y) \in \Delta \\ f(x,y) & (x,y) \notin \Delta \end{cases}. \tag{3}$$

Here, $\Theta(x,y)$ are the values of the known samples.

The convexity and closure of the above sets can be proved. Let us define

$$\mathcal{P} = \mathcal{P}_b \mathcal{P}_a, \tag{4}$$

The signal can be restored through the iteration

$$f^{(i+1)} = \mathcal{P} f^{(i)}. \tag{5}$$

# 3 Proposed semi-fragile watermarking methods

This section describes the proposed watermarking method which is an extension of the image content authentication method using the pinned sine transform (PST) proposed in [1]. The newly proposed watermarking method is discussed in Section 3.1. In Sections 3.2 and 3.3, we describe the processes of image authentication and restoration.

## 3.1 Watermark embedding

The watermark embedding process begins with dividing the original image into sub-blocks of size $n \times n$. These sub-blocks are then grouped into macro-blocks which contain $m \times m$ sub-blocks in raster order. Our authentication is based on sub-blocks while the restoration is based on macro-blocks. In the following, the words "block" and "sub-block" will be used alternatively.

Consider a macro-block $\mathbf{X}_\mu$, it is DCT transformed and the coefficients are re-ordered by zigzag scanning. The first $\ell m^2$ coefficients are kept and denoted as $\mathbf{d}_\mu$, where $\ell$ is the number of watermark bits to be embedded into every $n \times n$ sub-block. For example, $\ell = 6$, and $m = 3$, then the length of $\mathbf{d}_\mu$ is 54. The polarity information, $\mathbf{p}_\mu$, is generated by

$$p_\mu(k) = \begin{cases} 1 & d_\mu(k) \geq 0 \\ 0 & d_\mu(k) < 0 \end{cases}, \tag{6}$$

where $k = 0,\ 1,\ \ldots,\ \ell m^2 - 1$. The macro-blocks are then formed into pairs using a pre-determined mapping function $\Omega$. Suppose $\Omega(\mu) = \nu$, then $\mathbf{p}_\mu$, the polarity information of macro-block $\mathbf{X}_\mu$ is to be embedded into $\mathbf{X}_\nu$, and $\mathbf{p}_\nu$ is to be embedded into $\mathbf{X}_\mu$, with $\ell$ bits into each $n \times n$ block. A pseudo-random binary signal $\mathbf{h}$ is generated and its initial state number is contained as part of the secret key file $\mathcal{K}$. The watermark signal $\mathbf{w}_\mu$ is then obtained by XOR-ing the pseudo-random signal with the polarity information:

$$\mathbf{w}_\mu = \mathbf{h} \oplus \mathbf{p}_\mu. \tag{7}$$

The watermark is partitioned into $m^2$ parts, and each part is embedded into individual sub-blocks.

Consider an $n \times n$ block $\mathbf{x}$, it is first decomposed into two fields[1], the boundary field $\mathbf{x}^b$ and the pinned field $\mathbf{x}^p$, which can be described as:

$$\mathbf{x} = \mathbf{x}^b + \mathbf{x}^p. \tag{8}$$

Next, we perform the sine transform to the pinned field block as follows:

$$\mathbf{x}^{p(s)} = \mathbf{S}_n \mathbf{x}^p \mathbf{S}_n^T, \tag{9}$$

where $\mathbf{S}_n$ is the sine transform matrix of order $n$:

$$S_n(i,j) = \sqrt{\frac{2}{n+1}} \sin \frac{\pi(i+1)(j+1)}{n+1} \tag{10}$$

where $0 \leq i, j \leq n-1$.

In the middle to high frequency bands of $\mathbf{x}^{p(s)}$, we select, according to the length of the watermark signal, $\ell$ coefficients for watermarking modulation. A specific bit $w(k)$ is embedded into a coefficient $x^{p(s)}(k)$ according to the following algorithm:

**Algorithm 1** *Watermark embedding*

[1] Refer to [1] for the specific process.

***if*** $w(k) = 1$ ***then***
  ***if*** $x^{p(s)}(k) > \lambda$ ***then***
    $\hat{x}^{p(s)}(k) = \mathbf{x}^{p(s)}(k)$
  ***else***
    $\hat{x}^{p(s)}(k) = \alpha_1$
  ***end if***
***else if*** $w(k) = 0$ ***then***
  ***if*** $x^{p(s)}(k) < -\lambda$ ***then***
    $\hat{x}^{p(s)}(k) = x^{p(s)}(k)$
  ***else***
    $\hat{x}^{p(s)}(k) = \alpha_2$
  ***end if***
***end if***

where

- $\hat{x}^{p(s)}(k)$ is the corresponding watermarked coefficient;
- $\lambda$ is a sufficiently large threshold of positive value. It can be determined by users; its value will affect the tradeoff between the perceptual quality of the watermarked image and the robustness of the semi-fragile watermark;
- $\alpha_1$ and $\alpha_2$ are floating point values chosen randomly from $[\lambda/2,\ \lambda]$ and $[-\lambda, -\lambda/2]$, respectively.

The watermarked pinned field block $\hat{\mathbf{x}}^p$ is obtained by the inverse 2-D sine transform:

$$\hat{\mathbf{x}}^p = \mathbf{S}_n^T \hat{\mathbf{x}}^{p(s)} \mathbf{S}_n \tag{11}$$

and a watermarked block is therefore achieved by

$$\hat{\mathbf{x}} = \mathbf{x}^b + \hat{\mathbf{x}}^p. \tag{12}$$

### 3.2 Watermark Detection and Image Authentication

The The watermark detection and image authentication processes are described in this section. The detection system receives as input a watermarked and possibly tampered image. Similar to the watermarking process, the polarity information is extracted from every macro-block and is partitioned into $m^2$ parts, with every part corresponding to one sub-block in its paired macro-block. Here, we assume that the pre-determined mapping function $\Omega$ is known to both encoder and decoder. Since we limit ourselves to the situations in which tampering is only in the form of content modification, the synchronization issue after geometrical attacks is not considered here.

The embedded watermark is extracted from every sub-block by the following algorithm:

**Algorithm 2** *Watermark detection*

***if*** $\check{x}^{p(s)}(k) \geq 0$ ***then***
  $\check{w}(k) = 1$
***else***
  $\check{w}(k) = 0$
***end if***

where $\check{\mathbf{w}}$ is the extracted watermark. The extracted watermark is then XOR-ed with the corresponding part of the polarity information of its paired macro-block:

$$\check{\mathbf{h}} = \check{\mathbf{w}} \oplus \check{\mathbf{p}}. \tag{13}$$

The original pseudo-random signal $\mathbf{h}$ is also generated using the initial state number in $\mathcal{K}$. The bits in $\check{\mathbf{h}}$ and $\mathbf{h}$ are then compared by the normalized cross correlation function $\rho$, whose value lies in $[-1, 1]$. Assume $\gamma$ is a properly set threshold, the block is considered to be maliciously tampered if $\rho < \gamma$. The threshold is determined mathematically or experimentally so as to maximize the probability of tamper detection subject to a given probability of false alarm.

### 3.3 Content restoration

If some parts of the watermarked image were detected to be removed or destroyed, they would be automatically restored using the method described in Section 2. While our authentication is based on sub-blocks, the restoration is based on the macro-blocks. The macro-block containing tampered blocks is viewed as an irregular sampled signal with lost samples on the locations of the tampered blocks. The tampered blocks are then restored using the following algorithm:

**Algorithm 3** *Restoration of tampered blocks*

$$\begin{aligned} Init \quad & \mathbf{X}^{(0)} = \mathcal{P}_0 \mathbf{X} \\ & \mathbf{X}^{(i+1)} = \mathcal{P} \mathbf{X}^{(i)} \end{aligned}$$

where the projection operator $\mathcal{P}$ is as that defined in Eq. (4), and $\mathcal{P}_0$ is defined by

$$\mathcal{P}_0 \mathbf{y} = \begin{cases} 0 & \text{if } n \in \Delta \\ y(n) & \text{if } n \notin \Delta \end{cases} \tag{14}$$

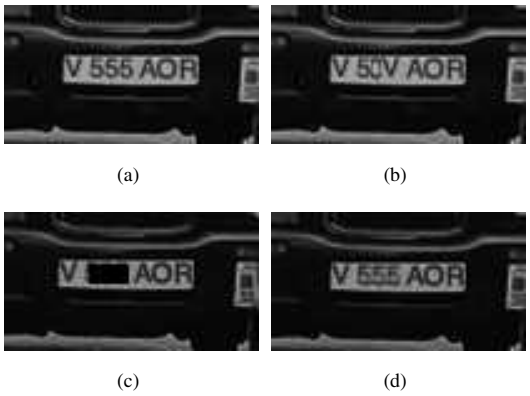
with $\Delta$ denoting the tampered sub-blocks. The polarity information, $\Gamma(u, v)$, has been extracted as described in the Section 3.2 from the paired macro-block of the tampered macro-block.

## 4 Simulation Results

The $256 \times 256$ traffic image as shown in Fig. 1(a)was used to test our algorithm. Figure 1(b) displays the watermarked image. We can see that the watermarked image looks identical to the original image, with PSNR of approximately 35 dB. In the watermarked image, the licence

(a) (b)

**Figure 1. Simulation results: (a) the original image; (b) the watermarked image.**



(a) (b)

(c) (d)

**Figure 2. Simulation results: (a) a portion of the watermarked image; (b) modified image; (c) authentication results and (d) restored result.**

number, as shown in full size in Fig. 2(a), was maliciously modified as illustrated in Fig. 2(b). The authentication and restoration results are shown in the Figs. 2(c) and 2(d), respectively. We observe that the tampered block was accurately detected and the restored result is acceptable with a PSNR of approximately 22 dB.

We also tested the robustness of our algorithm to various image processing operations and the results are tabulated in Table 1. The irregular sampling restoration process was tested using the same method as described in Section 3.3 assuming a null initial guess. It can be observed that our scheme maintains a low probability of $P_{FA} < 0.01$ for JPEG compression and other processing attacks including the median filtering, which is very threatening for the successful detection of watermarks. As for restoration, satisfactory results (e.g., PSNR of approximately 20 dB) are assured, except for the extreme situations of median filtering.

| Attacks | $P_{FA}$ | PSNR of Restored Blocks |
|---|---|---|
| *JPEG QF=30* | 0.0127 | 19.27 dB |
| *Gaussian filtering* | 0.0000 | 22.62 dB |
| *Unsharpening* | 0.0000 | 20.03 dB |
| *Contrast enhancement* | 0.0000 | 22.63 dB |
| *Salt & pepper noise* | 0.0054 | 20.87 dB |
| *Median filtering* | 0.0090 | 12.57 dB |

**Table 1. Robustness of authentication and self-restoration (*compared with the original unwatermarked image).**

## 5 Conclusion

In this paper, a semi-fragile watermarking method was proposed for automatic content authentication and restoration of traffic images. The problem of restoration of tampered images was expressed as an irregular sampling problem in discrete cosine transform subspace. The tampered image can be reconstructed through iterative projections onto convex sets. Prior knowledge was hashed into the watermark signal and embedded into the pinned field of PST of the original image, which contains the texture information of the original image. Experimental results showed that accurate authentication and restoration were still possible after lossy transcoding and other common image processing operations such as filtering.

## References


[1] A. T. S. Ho, X. Zhu, and Y. L. Guan. Image content authentication using pinned sine transform. *EURASIP Journal on Applied Signal Processing, Special Issue on Multimedia Security and Rights Management*, 2004(14):2174–2184, Oct. 2004.

[2] P. Marziliano and M. Vetterli. Irregular sampling in approximation subspaces. In *SampTA'99*, Loen, Norway, Aug. 1999.

[3] C. C. C. Pang, W. W. L. Lam, and N. H. C. Yung. A novel method for resolving vehicle occlusion in a monocular traffic-image sequence. *IEEE Trans. Intell. Transport. Syst.*, 5(3):129–141, Sept. 2004.

[4] C. Setchell and E. L. Dagless. Vision-based road-traffic monitoring sensor. In *IEE Proceedings on Vision, Image and Signal Processing*, pages 78–84, Feb. 2001.