# COLLUSION RESILIENT FINGERPRINT DESIGN BY ALTERNATING PROJECTIONS

*H. Oktay Altun, Gaurav Sharma, Adem Orsdemir, Mark F. Bocko*

ECE Dept., University of Rochester, Rochester, NY, USA

## ABSTRACT

Digital fingerprinting techniques aim to embed unique identification information into digital content distributed to individual users in order to track unauthorized use of multimedia files. Fraudulent users may not only attempt to remove the embedded signatures but also may form coalitions in order to remove the embedded fingerprint and disable tracking. This makes the design of fingerprints challenging. An effective fingerprint should not only carry the assigned users information but also guard against the possibility of falsely implicating an innocent user. Furthermore, in possible collusion scenarios, the colluded copies should identify each of the colluders. The embedded fingerprints should be imperceptible to maintain the commercial value of the content and preferably the fingerprint-based identification should survive content preserving signal processing. In this paper we give a precise description of each of these requirements and give a solution framework to obtain a set of fingerprinted images meeting these requirements.

***Index Terms—*** Digital fingerprinting, collusion resilience, POCS, adaptive marking.

## 1. INTRODUCTION

Today, digital multimedia dominates over conventional multimedia environments. Thanks to digital technology and the Internet, it is now easier to produce and sell digital files. However, the Internet also provides a non-centralized environment that is susceptible to unauthorized file sharing. This vulnerability causes significant financial losses in many sectors of the industry.

One of the effective ways of dissuading misuse of distributed digital multimedia copies is to hide a distinct secret message in each copy to identify the traitor. This method is known as fingerprinting. An important feature of fingerprinting technology compared to other data hiding techniques is its inherent design to track the colluding users who may collectively synthesize a copy for distribution from their individual copies. This makes the design of fingerprints challenging since the individual fingerprints embedded in the same content may be easily removed or attenuated during collusion.

There are several desirable qualities for a fingerprint design. Firstly, the fingerprint for each user must be detectable in the copy distributed to them. The interference between content and the embedded signal should be controlled in each copy. The endeavor will be void if the signal disappears within the content. The fingerprinted copy should not falsely indicate the presence of other fingerprints causing false accusations of innocent users. Secondly, the multimedia should keep its commercial or artistic value after the data hiding process. Visual quality degradation should be kept minimal by employing human visual system models, which allow degradations to be kept imperceptible.

Fingerprinting scenarios examined in the literature incorporate these requirements to varying degrees. The methods proposed include either coding for collusion resilience [1, 2] or design fingerprints in an orthogonal fashion to avoid interference between fingerprints [3, 4]. As an alternative we propose explicit mathematical modeling of the fingerprinting constraints. Fingerprinted images then can then be determined by set theoretic estimation methods. Specifically, we consider the design of spread spectrum fingerprints in this paper.

## 2. ANALYTICAL DESCRIPTIONS OF FINGERPRINTING REQUIREMENTS:

Figure 1 illustrates a typical fingerprinting scenario. The fingerprinted images, which look similar to the original copy, go through a channel where the copy can be compressed or noise may be added. Then the fingerprinted copies can be averaged to generate a colluded copy intended for fraudulent distribution. In order to generate fingerprinted copies that satisfy all these requirements is challenging. We first mathematically model all these requirements and give a mathematical framework that satisfy all these requirements simultaneously.

Next we examine basic requirements of fingerprinting in the form of constraint sets. A typical set of fingerprinted images should satisfy these requirements simultaneously. For simplicity we will consider constraints on individual images. These extend to the entire collection thorough a simple product space formulation. A spread spectrum technique is employed to embed the fingerprint information into the sequence.

### 2.1. Identification of individual traitors:

By correlating the suspicious copy with each assigned pseudo-random sequence a colluding user may be identified with high confidence. In general, each fingerprinted copy should give a positive response with key generated pseudo-random sequence. The number of sets in this category is equal to the number of distributed copies (K).

$$S_1^j \equiv \{\mathbf{x}_j : (\mathbf{w}_j - \overline{\mathbf{w}}_j)^T(\mathbf{x}_j - \overline{\mathbf{x}_j}) \geq \tau_a\}, j = 1, \ldots, K. \quad (1)$$

$$\equiv \{\mathbf{x}_j : \mathbf{w_j}^T(\mathbf{x}_j - \overline{\mathbf{x}_j}) \geq \tau_a\},\ j = 1, \ldots, K. \quad (2)$$

We denote the $j^{th}$ fingerprinted image by $x_j$, the complete collection of the fingerprinted images by $\mathbf{x} = [\mathbf{x}_1^T \ldots \mathbf{x}_K^T]$ and j'th user pseudo-random sequence generated using a corresponding key by $w_j$. An image size pn-sequence is correlated with mean correction and compared against a threshold for detection [5]:

The threshold $\tau_a$ is considered to be same for all agents in the present paper. However, in situations where different groups of users may present different risk levels, these values can be set according to the corresponding risk of the groups.

This work was supported by the the Air Force Office of Scientific Research under grant number FA9550-07-1-0017.

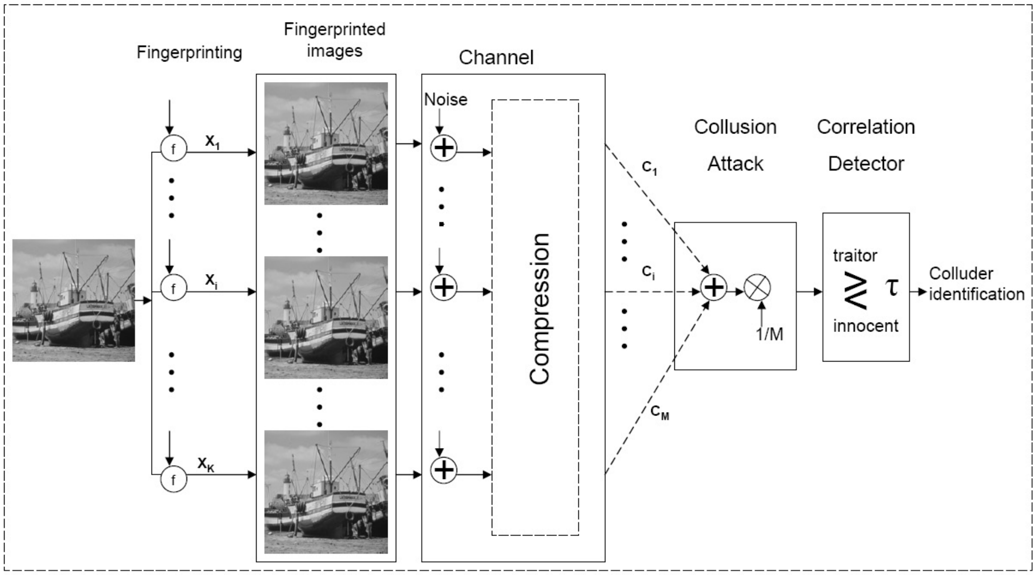

**Fig. 1**. Fingerprinting scenarios.

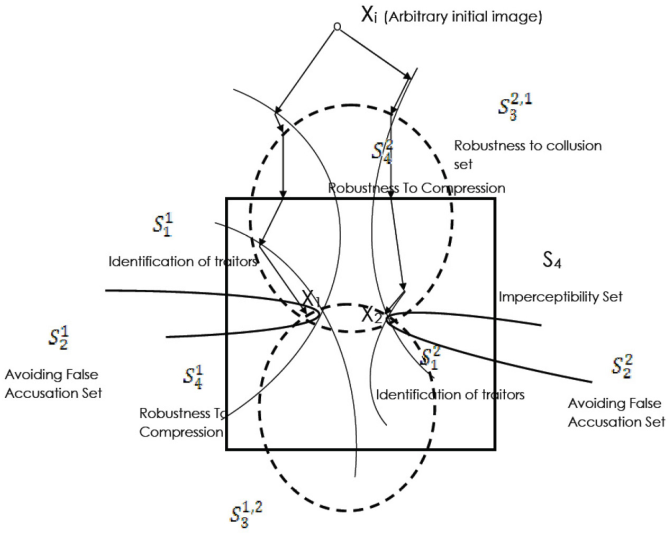
Xi (Arbitrary initial image)
$S_3^{2,1}$
$S_4^2$
Robustness to collusion set
Robustness To Compression
$S_1^1$
Identification of traitors
$S_4$
Imperceptibility Set
X1
X2
$S_2^1$
Avoiding False Accusation Set
$S_4^1$
Robustness To Compression
$S_1^2$
Identification of traitors
$S_2^2$
Avoiding False Accusation Set
$S_3^{1,2}$

**Fig. 2**. Generic illustration of constraint sets and obtaining fingerprinted images in a 2-colluder scenario.

### 2.2. Avoiding accusation of benign individual:

The innocent user's secret sequence should not correlate positively with other people's copies. Mathematically speaking, the cross-correlation between user's pn-sequences and copies should be kept below some threshold value.

$$S_2^{i,j} \equiv \{\mathbf{x_j} : \mathbf{w}_i^T(\mathbf{x}_j - \overline{\mathbf{x}_j}) \leq \tau_b\}, \;\; i,j = 1, \ldots, \; K, i \neq j. \quad (3)$$

There are $K \times (K-1)$ possible crossings between keys and copies. This is the number of constraint sets to avoid accusation of innocent users.

### 2.3. Identification of each colluder in all possible collusion scenarios:

There are different strategies for identification of colluders. The possible goals of fingerprinting can be listed as: Catch-all, catch-some and catch-one strategy [3]. We consider catch-all strategy and only the linear averaging attack of all possible coalitions. The correlation detector should identify each colluder participating in the collusion. This requirement can be mathematically described as follows:

$$S_3^{i,k} \equiv \{\mathbf{x_j} : \mathbf{w}_i^T(\widehat{\mathbf{x}}_{C_k} - \overline{\widehat{\mathbf{x}}_{C_k}}) \geq \tau_a\}, \;\; \forall (k, i \in C_k). \quad (4)$$

where $C_k$ represents all possible coalition scenarios. There are $2^K - K - 1$ sets corresponding to all the possible collusion scenarios.

### 2.4. Imperceptibility of fingerprints:

Each fingerprinted copy must maintain fidelity to the original image. In order to ensure this we incorporate an imperceptibility constraint for the perturbation introduced by fingerprints. Particularly, we employ a spatial domain texture masking model proposed by Voloshynovsky et al. [6]. The model provides pixel-wise upper and lower bounds for the difference from the original image. The resulting constraint can be expressed as:

$$S_4^j \equiv \{\mathbf{x_j} : \mathbf{l} < \mathbf{x_j} - \mathbf{x_0} < \mathbf{u}, \;\; \mathbf{j} = \mathbf{1}, \ldots, \; \mathbf{K}\}. \quad (5)$$

where $\mathbf{u}$ and $\mathbf{l}$ form pixel-wise upper and lower bounds respectively and $\mathbf{x}_o$ is the original image.

### 2.5. Robustness to compression:

Robustness against content preserving signal processing is desirable. Here, we concentrate on a common type of non-malicious signal processing operations: lossy compression, specifically JPEG compression. JPEG compression is performed by quantization of DCT coefficients of an image at a predetermined rate based on the desired image quality. The spread spectrum watermark is robust against compression if the detector response is above threshold after compression. The set of images that illustrates this kind of robustness can be approximated as [7]:

$$\widehat{S}_5^j \equiv \{\mathbf{x}_j : \mathbf{w_j}^T (IDCT(Q_0[DCT(\mathbf{x}_j)] - \overline{IDCT(Q_0[DCT(\mathbf{x}_j)])} \geq \gamma\},\ j = 1, \ldots, K. \quad (6)$$

$$Q_0^k[t] = \begin{cases} 0 & Q^k[(DCT(\mathbf{x}_0))_k] = 0 \\ t & \text{otherwise} \end{cases} \quad ; \quad k = 0, 1, \ldots, MN - 1. \quad (7)$$

where $(DCT(\mathbf{x}_0))_k$ denotes the $k^{th}$ transform coefficient of the original image $\mathbf{x}_0$ and $MN$ denotes the size of the image. Thus the quantizer $Q_0[\ ]$ sets the transform coefficients that are zero in $Q[DCT(\mathbf{x}_0)]$ to zero and leaves other coefficients unchanged providing a subspace projection approximations to JPEG compression.

### 2.6. Robustness against Gaussian noise attacks:

Many alterations on the multimedia can be modeled as Gaussian noise. Hence, achievement of robustness against noise is important. It is hard to express this set analytically as in robustness to compression set however, it can be handled by asymmetry between the embedding and detection thresholds. The embedder may aim $\tau_a$ in embedding but the detector can compare with $\tau_a - \Delta$, where $\Delta$ can be chosen depending on the expected noise on the media.

## 3. PROPOSED METHOD AND EXPERIMENTAL RESULTS

The proposed constraints sets for the fingerprinting requirements are all convex sets. The sets corresponding to identification of individual traitors, avoiding accusation of benign individuals, identification of each colluder in all possible collision scenarios, imperceptibility of fingerprint, and robustness to compression sets are all affine [7]. A set of images that satisfies all these requirements simultaneously can be found by the iterative algorithm of successive projections onto convex sets(POCS) [8]. Given $n$ convex sets $\{S_i\}_{i=1}^n$ the POCS method determines a point in their intersection by successive projections. If the intersection set is non-empty, the sequence $\{f_k\}_{k=0}^\infty$ generated by successive projections onto the sets converges to a point in the intersection, where

$$f_{k+1} = (P_{S_n}(P_{S_{n-1}} ... P_{S_1}(f_k)...)), k = 0, 1, .. \quad (8)$$

where $P_{S_i}$ is the projection operator onto set $S_i$, defined as $P_{S_i} = arg\ min_{y \in S_i} \|y - x\|$.

Figure 2 illustrates the constraints in a 2-colluder scenario where for the purpose of illustration we consider the constraints on the individual images (instead of collection together). $x_i$ represents an arbitrary initial image. There are two different projection paths, leading to the two distinct fingerprinted images. The sets represented by dashed lines are "robustness to collusion" sets and effective for both projection sequences. The square shaped imperceptibility set is also common to both fingerprinted image generation processes. The rest of the sets have distinct counterparts for each projection sequence. For example, avoiding false accusation sets are distinct and work independently from each other. Although, Fig. 2 gives an explanation of the formation of distinct sequences in the fingerprinting process, the dynamics between the images need further elaboration.

We examined the algorithm on 8 different 512x512 images from the USC database. The number of agents is constrained to be 3, to keep the number of constraints low. All sets except robustness to compression set are employed. The values $\tau_a$ and $\tau_b$ are chosen to be 2 and 0, respectively. The $S_0$ and $S_1$ parameters of the noise visibility texture model (see [6] for details) are chosen to be 20 and 2, respectively.

The Boat image, fingerprinted version and error image (difference between watermarked and original) are illustrated in Figures 3,4, 5 respectively. The difference image is scaled in order to make it readily perceptible and shifted to a midgray average value so as to accommodate negative values. The difference image shows the adaptation of the embedded information onto the original image. The PSNR values of three resulting images are 31.64 dB, 31.62 dB and 32.19 dB.($PSNR_j = N^2 225^2 / (\sum_i (\mathbf{x}_o(i) - \mathbf{x}_j(i))^2)$ where $\mathbf{x}_j$ is fingerprinted image and $\mathbf{x}_o$ is the original image).

In this particular scenario we did not explicitly perform the detection at the receiver. Instead we provide the correlation values at the receiver (that form the input to the threshold based detector). Particularly we have set $\tau_a = 2$ and $\tau_b = 0$. The $\Delta$ parameter can be set adaptively according to the expected attacks on the multimedia. For instance, choosing $\Delta = 0.5$ would set the detector threshold to a correlation value of $1.5$. This would allow the colluded copies some noise margin in order to accommodate other types of attacks.

We next demonstrate that the correct keys identify the correct images and possible colluded images. The only collusion strategy considered is linear averaging attacks. The responses of each key with the 3 different fingerprinted images and all possible combinations of attacks are tabulated in Table 1 . Table 2 illustrates the average statistics over the 8 different images. All entries correspond to mean-corrected correlation values. We can see that the correlation values observed for all individual and colluded copies allow for the identification of colluders while avoiding accusations of innocent individuals.

Due to the perceptual shaping for the same cover, the fingerprints are dependent upon the cover and thereby upon each other. This can be explicitly seen in the difference image in Fig. 5. Thus the embedded watermark signals cannot implicitly be assumed to be orthogonal to each other. However, by virtue of the constraints corresponding to benign users, in the set of fingerprinted images, each image has a zero cross-correlation with other users' pn-random fingerprint sequences. This is more meaningful in typical oblivious detection scenarios.

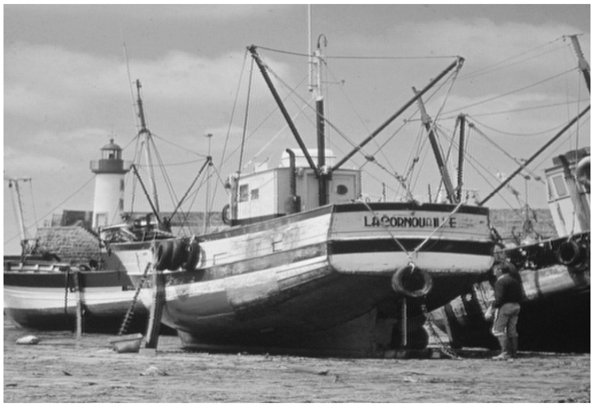

**Fig. 3**. Original image.

|  | $Image$ 1 | $Image$ 2 | $Image$ 3 | $Average$ (1&2) | $Average$ (1&3) | $Average$ (2&3) | $Average$ (1&2&3) |
|---|---|---|---|---|---|---|---|
| $Key$ 1 | 5.989 | 0.004 | 0.004 | 2.997 | 2.997 | 0.004 | 1.999 |
| $Key$ 2 | 0.004 | 5.99 | 0.004 | 2.997 | 0.004 | 2.997 | 1.999 |
| $Key$ 3 | 0.003 | 0.003 | 5.993 | 0.0036 | 2.898 | 2.998 | 2.0 |

**Table 1**. Correlation values between the key controlled pn-sequences and associated fingerprinted images for boat image.

|  | $Image$ 1 | $Image$ 2 | $Image$ 3 | $Average$ (1&2) | $Average$ (1&3) | $Average$ (2&3) | $Average$ (1&2&3) |
|---|---|---|---|---|---|---|---|
| $Key$ 1 | 5.984 | 0.007 | 0.007 | 2.995 | 2.995 | 0.007 | 1.999 |
| $Key$ 2 | 0.007 | 5.983 | 0.007 | 2.995 | 0.007 | 2.995 | 1.999 |
| $Key$ 3 | 0.006 | 0.007 | 5.986 | 0.007 | 2.996 | 2.996 | 2.0 |

**Table 2**. Average of correlation values between the key controlled pn-sequences and associated fingerprinted images for 8 USC images.

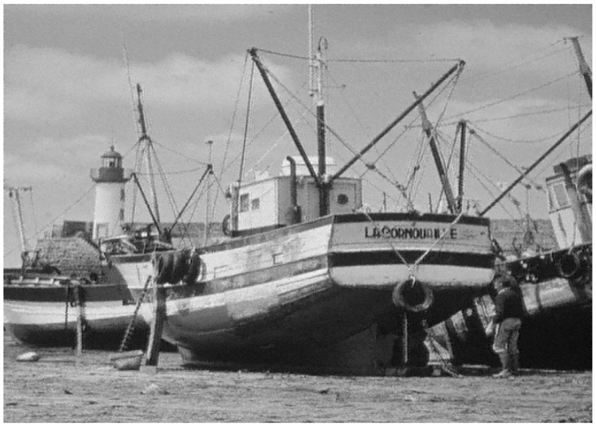

**Fig. 4**. Fingerprinted image. PSNR is 31.64 dB

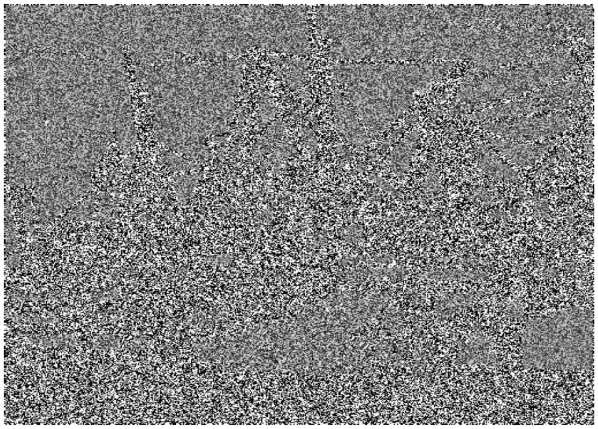

**Fig. 5**. Difference image between fingerprinted image 1 and original image.

## 4. CONCLUSION

Fingerprinting requirements may be formulated as constraint sets allowing a set theoretic estimation framework to be utilized to determine fingerprinted images. Many common requirements such as identification of traitors, false accusation probability, identification in the presence of collusion and compression, can be formulated as convex constraints allowing a solution to the problem using the method of projections onto convex sets (POCS).

We described fingerprinting requirements as convex constraints and achieved fingerprinted copies by applying set theoretical framework. As a proof of concept we designed a 3-agent fingerprinted copies of an image. One drawback of the framework is the rapidly increasing number of constraints due to vast amount of possible combination of coalitions. However, the technique can be appropriate for scenarios where there are a small number of fingerprinted copies(e.g Hollywood screeners).

## 5. REFERENCES

[1] Dan Boneh and James Shaw, “Collusion-secure fingerprinting for digital data,” *Lecture Notes in Computer Science*, vol. 963, pp. 452–464, 1995.

[2] Y. Yacobi, “Improved boneh-shaw content fingerprinting,” in *Proc. CTRSA*, 2001, pp. 378–391.

[3] M.Wu, W.Trappe, Z.Wang, and K.J.R.Liu, “Collusion-resistant fingerprinting for multimedia,” *IEEE Signal Process. Magazine*, vol. 21, pp. 15–27, Mar 2004.

[4] J. Kilian, T. Leighton, L. Matheson, T. Shamoon, R. Tarjan, and F. Zane, “Resistance of digital watermarks to collusive attacks,” in *Proc. IEEE Int.Symp. Inform. Theory*, Aug 1998, p. 271.

[5] I. Cox, J.Kilian, F.T.Leighton, and T.Shamoon, “Secure spread spectrum watermarking for multimedia,” *IEEE Trans. Image Proc.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[6] Sviatoslav Voloshynovskiy, Alexander Herrigel, Nazanin Baumgaertner, and Thierry Pun, “A stochastic approach to content adaptive digital image watermarking,” in *Information Hiding*, 1999, pp. 211–236.

[7] Oktay Altun, Gaurav Sharma, Mehmet Celik, and Mark Bocko, “A set theoretic framework for watermarking and its application to semifragile tamper detection,” *IEEE Trans. Info. Forensics and Security*, vol. 1, no. 4, pp. 479–492, December 2006.

[8] P. L. Combettes, “The foundations of set theoretic estimation,” *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, Feb. 1993.