Semi-Fragile Watermarking in Biometric Systems: Template Self-Embedding *

Reinhard Huber[1], Herbert Stögner[1], and Andreas Uhl[1,2]

[1] School of CEIT, Carinthia University of Applied Sciences, Austria
[2] Department of Computer Sciences, University of Salzburg, Austria

Contact author e-mail: uhl@cosy.sbg.ac.at

Abstract. Embedding biometric templates as image-dependent watermark information in semi-fragile watermark embedding is proposed. Experiments in an iris recognition environment show that the embedded templates can be used to verify sample data integrity and may serve additionally to increase robustness in the biometric recognition process.

1 Introduction

There has been a lot of work done during the last years proposing watermarking techniques to enhance biometric systems security in some way (see [4] for our recent survey on the topic). Major application scenarios include biometric watermarking (where biometric templates are embedded as "message" as opposed to classical copyright information), sample data replay prevention (by robustly watermarking once acquired sample data), covert biometric data communication (by steganographic techniques), and employing WM is a means of tightly coupled transport of sample data and embedded (template or general purpose authentication) data for multibiometric or two-factor authentication schemes, respectively.

In this work we consider the application scenario where the aim of WM is to ensure the integrity and authenticity of the sample data acquisition and transmission process. During data acquisition, the sensor (i.e. camera) embeds a watermark into the acquired sample image before transmitting it to the feature extraction module. The feature extraction module only proceeds with its tasks if the WM can be extracted correctly (which means that (a) the data has not been tampered with and (b) the origin of the data is the correct sensor).

Attack An attacker aims at *inserting* the WM in order to mimic correctly acquired sensor data or to *manipulate* sample data without affecting the WM.

WM properties and content The WM needs to be unique in the sense that it has to uniquely identify the sensor. Resistance against a WM insertion attack can be achieved by sensor-key dependent embedding. Since the watermarking

* This work has been partially supported by the Austrian Science Fund, project no. L554-N15.

scheme has to be able to detect image manipulations, (semi-)fragile embedding techniques are the method of choice. Especially in semi-fragile watermarking it was found to be highly advantageous to embed image-dependent watermark data in order to prevent copy attacks. WM extraction should be blind.

Crypto alternative Classical authentication protocols can be used to secure the communication between sensor and feature extraction module – a digital signature signed with the private key of the acquisition device can ensure the authenticity of the sensor and the integrity of the image data. However, this approach cannot provide robustness and no information about tampering locations is obtained.

Yeung et al. [9] propose a fragile watermarking technique to add the ability for integrity verification of the captured fingerprint images against altering during transmission or in a database. Ratha et al. [8] propose to embed a response to an authentication challenge sent out by a server into a WSQ compressed fingerprint image in order to authenticate the sensor capturing the fingerprint image. If the (fragile) watermark cannot be extracted, either the image has been tampered with or the image does not come from the correct sensing device.

Also, semi-fragile watermarking has been suggested to verify authenticity of biometric sample data. PCA features are used as embedded data in [7], while [1] proposes the embedding of robust signatures into fingerprint images.

Finally, dual WM techniques have been proposed applying two different embedding techniques concurrently. The first technique in [6] is used for checking integrity on a block level using CRC checks, the second provides reversible watermarking in case the first technique rates the sample as being authentic. Two different embedding techniques (a semi-fragile and a robust one) for embedding both, a sample image dependent signature as well as a template of a different modality are proposed by Komninos et al. [5].

In this paper we focus on protecting the transmission of sample data from the sensor to the feature extraction module employing a specific semi-fragile watermarking technique. In particular, we propose to embed biometric template data instead of general purpose watermark information which can then be used in the matching process in addition to checking integrity. In Section 2, we introduce the template-embedding based semi-fragile watermarking approach and discuss its properties. Section 3 presents experiments where the proposed concept is evaluated in the context of iris recognition using a variant of a well known watermark embedding scheme. Section 4 concludes the paper.

2 Semi-Fragile Watermarking by Template Self Embedding

In the context of biometrics, we propose to embed template data as semi-fragile WM information instead of general purpose image descriptors as used in classical semi-fragile WM schemes [2]. This is sensible since on the one hand template data are of course image dependent data and therefore are able to prevent WM copy

attacks or similar. On the other hand, in case of tampering or other significant data manipulations, the aim is not to reconstruct the sample data at first hand, but to be able to generate template data from the sample data required for matching. So data for reconstructing the sample data is suggested to be replaced by data for directly generating template data. In the following, we describe the WM embedding and extraction processes:

1. From the acquired sample data, a template is extracted.
2. The template is embedded into the sample data employing a semi-fragile embedding technique (this template is referred to as "template watermark" subsequently).
3. The data is sent to the feature extraction and matching module.
4. At the feature extraction module, the template watermark template is extracted, and is compared to the template extracted from the sample (denoted simply as "template" in the following). In this way, the integrity of the transmitted sample data is ensured when there is sufficient correspondence between the two templates. In case of a biometric system operating in verification mode the template watermark can also be compared to the template in the database corresponding to the claimed identity (denoted "database template" in the following).
5. Finally, in case the integrity of the data has been proven, the watermark template and the template are used in the matching process, granting access if the similarity to the database template(s) is high enough.

When comparing this approach to previous techniques proposed in literature, we notice the following differences / advantages: As opposed to techniques employing robust template embedding watermarking (e.g. as proposed for enabling tightly coupled transport of sample and template data of different modalities), the proposed scheme can ensure sample data integrity. The importance of this property has been recently demonstrated [3] in an attack against robust embedding schemes used in the multibiometric and two-factor authentication scenarios. As opposed to techniques employing arbitrary (semi-)fragile watermarks for integrity protection (instead of the template watermark used here), the template watermark data can be used to provide a more robust matching process after data integrity has been assured.

However, some issues need to be investigated with respect to the proposed scheme (which will be done in the experiments):

- Does integrity verification indeed work in a robust manner ?
- What is the impact of the embedded template watermark on the recognition performance using the template for matching only ?
- Can a combination of template watermark and template result in more robustness in an actual matching process ?

3 Experiments in the Case of Iris Recognition

3.1 Iris Recognition and Iris Databases

The employed iris recognition system is Libor Masek's Matlab implementation[3] of a 1-D version of the Daugman iris recognition algorithm. First, this algorithm segments the eye image into the iris and the remainder of the image. Iris image texture is mapped to polar coordinates resulting in a rectangular patch which is denoted "polar image". For feature extraction, a row-wise convolution with a complex Log-Gabor filter is performed on the polar image pixels. The phase angle of the resulting complex value for each pixel is discretized into 2 bits. The 2 bit of phase information are used to generate a binary code. After extracting the features of the iris, considering translation, rotations, and disturbed regions in the iris (a noise mask is generated), the algorithm outputs the similarity score by giving the Hamming distance between two extracted templates.

The following three datasets are used in the experiments:

CASIAv3 Interval database[4] consists of 2639 images with 320×280 pixels in 8 bit grayscale .jpeg format, out of which 500 images have been used in the experiments.

MMU database[5] consists of 450 images with 320×240 pixels in 24 bit grayscale .bmp format, all images have been used in the experiments.

UBIRIS database[6] consists of 1876 images with 200×150 pixels in 24 bit colour .jpeg format, out of which 318 images have been used in the experiments.

3.2 The Watermarking Scheme

As the baseline system, we employ the fragile watermarking scheme as developed by Yeung et. al and investigated in the context of fingerprint recognition [9]. For this algorithm, the watermark embedded is binary and padded to the size of the host image. Subsequently, the WM is embedded into each pixel according to some key information. As a consequence, the WM capacity is 89600, 76800, and 30000 bits for CASIAv3, MMU, and UBIRIS, respectively.

Since this technique is a fragile WM scheme, no robustness against any image manipulations can be expected of course. However, the usually smaller size of biometric templates can be exploited to embed the template in redundant manner, i.e. we embed the template several times. After the extraction process, all template watermarks are used in a majority voting scheme which constructs a "master" template watermark. We expect to result in higher robustness as compared to the original algorithm due to redundant embedding leading to an overall quasi semi-fragile WM scheme for the watermark templates. In our implementation, the iris code consists of 9600 bits, therefore, we can embed 9, 8, and 3 templates into images from the CASIAv3, MMU, and UBIRIS databases, respectively.

[3] `http://www.csse.uwa.edu.au/~pk/studentprojects/libor/sourcecode.html`
[4] `http://www.cbsr.ia.ac.cn/IrisDatabase.htm/`
[5] `http://pesona.mmu.edu.my/~ccteo/`
[6] `http://www.di.ubi.pt/~hugomcp/investigacao.htm`

Note that instead of this embedding scheme, any semi-fragile WM scheme [2] with sufficient capacity to embed template information can be employed.

3.3 Experimental Results

As first topic, we investigate integrity verification under conditions which require robustness properties. As "attacks" against the sample data with embedded WM, we consider mean filtering, noise addition, and JPEG compression. As a first scenario S1 (restricted to the verification scenario), comparison between extracted template WM and database (DB) template is covered. We consider the case that 5 different templates are stored in the database out of which a single database template is generated by majority coding like explained before in the case of the template WM. Table 1 (left) shows the bit error rate (BER) for the different attacks considered. The second scenario S2 is the comparison between extracted template WM and the template extracted from the watermarked sample data the results of which are shown in Table 1 (right).

	DB template vs. template			template WM vs. template		
Attack	CASIAv3	MMU	UBIRIS	CASIAv3	MMU	UBIRIS
No attack	0.21	0.23	0.19	0.14	0.06	0.07
Mean filtering	0.49	0.50	0.50	0.49	0.50	0.50
Gaussian Noise $N = 0.0005$	0.21	0.23	0.19	0.14	0.06	0.07
Gaussian Noise $N = 0.001$	0.21	0.23	0.19	0.14	0.06	0.07
JPEG Q100	0.21	0.23	0.19	0.14	0.06	0.08
JPEG Q99	0.21	0.24	0.22	0.14	0.07	0.11
JPEG Q98	0.25	0.30	0.32	0.20	0.18	0.26
JPEG Q95	0.41	0.45	0.45	0.39	0.41	0.44

Table 1. BER for seven different attacks.

The first thing to note is that even without attack, BER is clearly above zero. For S2 this effect is solely due to the influence the embedded WM has on the extracted template - obviously the WM changes the sample in a way that about 10% of the bits are altered. For S1 the differences are higher which is clear since the DB template is constructed from several distinct templates. We have to consider that a typical decision threshold value for the iris recognition system in use is at a BER in $[0.3, 0.35]$. When taking this into account, the extent of template similarity is of course enough to decide on proven sample integrity. For both S1 and S2, adding noise and applying JPEG compression with quality set to 100 (Q100) does not change the BER. When decreasing JPEG quality to 98, BER starts to increase slightly. The situation changes drastically when applying JPEG Q95 and mean filtering: BER is up to 0.4 - 0.5 which means that integrity cannot be verified successfully. We realize that integrity verification in our technique is indeed robust against moderate JPEG compression and noise. On the other hand, mean filtering and JPEG compression at quality 95% destroys the template WM and indicates modification. The distribution of incorrect bits can be used to differentiate between malicious attacks (where an accumulation of in-

correct bits can be observed in certain regions) and significant global distortions like compression where incorrect bits are spread across the entire data.

S1 and S2 can be combined into a single integrity verification scheme. The idea is to combine the single templates extracted from the watermark and the template extracted from the watermarked sample into a weighted "fused template": in our example, we use 4 copies of the template and the embedded number of templates from the template WM in a majority voting scheme to generate the fused template. Table 2 shows the corresponding BER when comparing the fused template to the DB template.

Attack	CASIAv3	MMU	UBIRIS
No attack	0.21	0.21	0.21
Mean filtering	0.30	0.27	0.21
Gaussian Noise $N = 0.0005$	0.21	0.21	0.21
Gaussian Noise $N = 0.001$	0.21	0.21	0.21
JPEG Q100	0.21	0.21	0.21
JPEG Q99	0.21	0.21	0.21
JPEG Q98	0.23	0.23	0.21
JPEG Q95	0.27	0.26	0.21

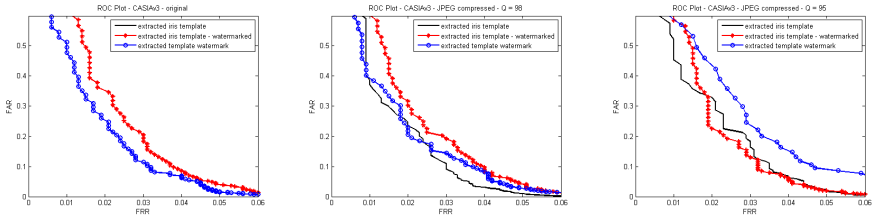
Table 2. BER for the fused template under seven different attacks.

It can be clearly seen that while the BER without attack and applying moderate attacks is higher as compared to S2, we get much better robustness against JPEG Q95 and even mean filtering. With the fusing strategy, robustness even against those two types of attacks can be obtained. Of course, the fusion scheme does only make sense in a biometric system in verification mode, since integrity verification is done against templates stored in the template database.

As a second topic, we investigate iris recognition performance using the template extracted from the watermarked sample (W1) and the extracted template WM (W2), and compare the behavior to the "original" results using templates extracted from the original sample data (without embedded WM, W0). For this purpose, we compare ROC curves of the three cases with and without attacks (i.e. JPEG compression, noise insertion, and mean filtering) conducted against the sample data.

In both Figs. 1.a and 2.a the curve W0 is hidden by W2 and we clearly note that the embedded WM impacts on recognition performance since W1 shows clearly inferior ROC (note that this contrasts to the case of fingerprint matching reported in [9]). So without attack, using the template WM is beneficial over the template. This situation is also typical for moderate attacks being conducted as shown in Figs. 1.b and 2.b as an example for the case of JPEG compression with Q98. While for the CASIAv3 data W0 and W2 are close, both being superior to W1, for the UBIRIS data W2 is the best option. W0 is clearly inferior to W2, while W1 is the worst option. Obviously, the embedded template watermark is not yet severely impacted by the compression artifacts.

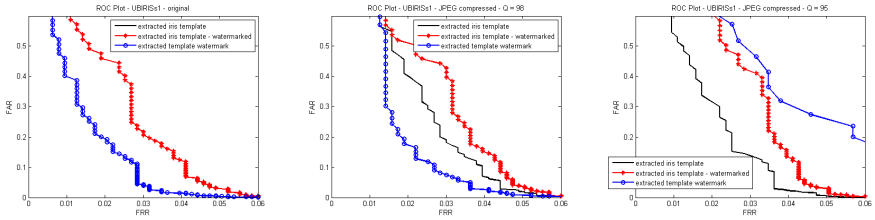
The situation changes when the attacks get more severe. As shown in Figs. 1.c and 2.c, under JPEG compression with Q95 W2 is the worst option now since



(a) no attack (b) JPEG Q98 (c) JPEG Q95

Fig. 1. ROC curves of the CASIAv3 data.

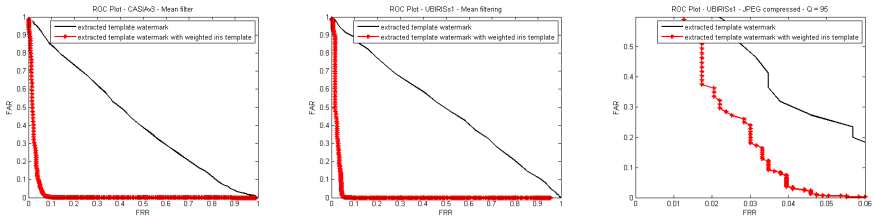
the robustness of the WM is not sufficient any more. While for the CASIAv3 data W0 and W1 are close (so the impact of the WM is negligible), for UBIRIS the impact of the WM is quite significant (which can be explained by the fact that the UBIRIS data is of already quite low quality without any further degradation, the additional WM complicates template extraction). For mean filtering the result for W2 is even worse as shown in Figs. 3.a and 3.b, no recognition can be performed at all with the extracted template WM after this attack.



(a) no attack (b) JPEG Q98 (c) JPEG Q95

Fig. 2. ROC curves of the UBIRIS data.

Finally, the strategy of combining W1 and W2 into a fused template for integrity verification (results given in Table 2) can also be applied for matching. Fig. 3 shows examples where the ROC behavior of W2 can be significantly improved by using this approach. In particular, in the case of mean filtering the fused template can be used for recognition purposes as shown in Figs. 3.a and 3.b.



(a) CASIAv3, mean (b) UBIRIS, mean (c) UBIRIS, Q95

Fig. 3. ROC curves for fused templates.

4 Conclusion

We have introduced the concept of embedding biometric templates as image-dependent watermark information in semi-fragile watermark embedding which serves the purpose of verifying the integrity and authenticity of the sensor - feature extraction communication. Experiments in an iris recognition environment show the feasibility of the approach and demonstrate, that the embedded templates can be used to verify integrity and may serve additionally as a means to increase robustness in the biometric recognition process.

References

[1] F. Ahmed and I.S. Moskowitz. Composite signature based watermarking for fingerprint authentication. In *Proceedings of the ACM Workshop on Multimedia and Security (MMSEC'05)*, pages 799 – 802, 2005.

[2] Ö. Ekici, B. Sankur, and M. Akcay. A comparative evaluation of semi-fragile watermarking algorithms. *Journal of Electronic Imaging*, 13(1):209–216, 2003.

[3] J. Hämmerle-Uhl, K. Raab, and A. Uhl. Attack against robust watermarking-based multimodal biometric recognition systems. In C. Vielhauer et al., editor, *Proceedings of the 2011 BioID Workshop*, volume 6583 of *Springer LNCS*, pages 25–36, Brandenburg, Germany, 2011.

[4] J. Hämmerle-Uhl, K. Raab, and A. Uhl. Watermarking as a means to enhance biometric systems: A critical survey. In A. Ker, S. Craver, and T. Filler, editors, *Proceedings of the 2011 Information Hiding Conference (IH'11)*, Springer LNCS, Prague, Czech Republic, 2011. to appear.

[5] N. Komninos and T. Dimitriou. Protecting biometric templates with image watermarking techniques. In *Advances in Biometrics (Proceedings of ICB 2007)*, volume 4642 of *Springer Lecture Notes on Computer Science*, pages 114 – 123, 2007.

[6] H. Lee, J. Lim, S. Yu, S. Kim, and S. Lee. Biometric image authentication using watermarking. In *Proceedings of the International Joint Conference SICE-ICASE, 2006*, pages 3950 – 3953, 2006.

[7] C. Li, B. Ma, Y. Wang, and Z. Zhang. Protecting biometric templates using authentication watermarking. In *Advances in Multimedia Information Processing - PCM 2010*, volume 6297 of *Springer Lecture Notes on Computer Science*, pages 709 – 718, 2010.

[8] Nalini K. Ratha, Miguel A. Figueroa-Villanueva, Jonathan H. Connell, and Ruud M. Bolle. A secure protocol for data hiding in compressed fingerprint images. In *Proceedings of the BioAW 2004*, volume 3087 of *Lecture Notes in Computer Science*, pages 205–216, 2004.

[9] Minerva M. Yeung and Sharat Pankanti. Verification watermarks on fingerprint recognition and retrieval. *Journal of Electronal Imaging, Special Issue on Image Security and Digital Watermarking*, 9(4):468–476, October 2000.