Two-Factor Biometric Recognition with Integrated Tamper-protection Watermarking

Andreas Uhl

Department of Computer Sciences
University of Salzburg, Austria

uhl@cosy.sbg.ac.at

http://www.wavelab.at/

Outline

- Introduction & Motivation
- Watermarking in Biometrics
- Two-factor authentication Approach
- Experiments
- Conclusion

Introduction

With the increasing popularity of biometric recognition applications, several security breaches have been discovered. Watermarking (WM) has been suggested as a means to resolve some of these problems as well as to add additional functionalities to biometric systems.

We address a two-factor authentication system, where data stored on a smart-card is embedded into biometric sample data by means of a **semi-fragile** watermarking scheme. The smart-card data consists of a biometric template of the same modality as the sample data thus resulting in a multibiometric recognition scheme with eventually improved recognition performance and additional features.

Watermarking in Biometrics

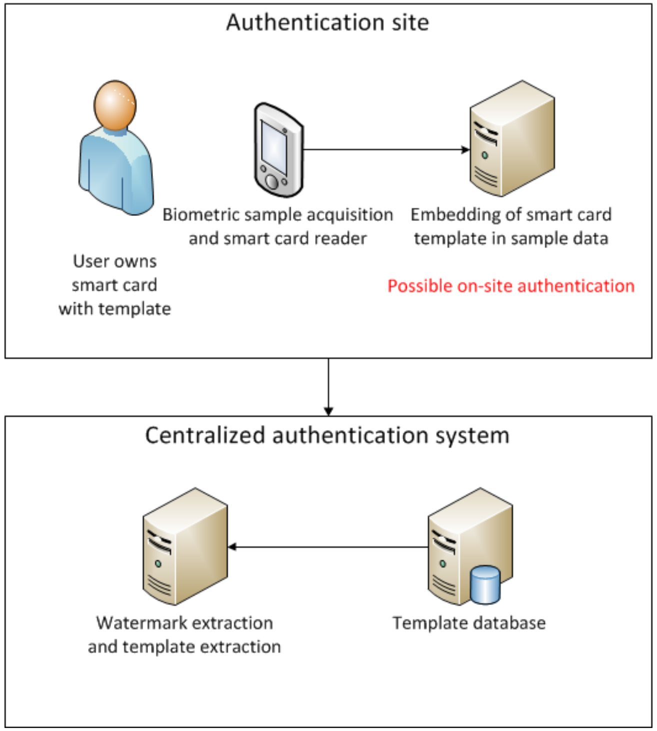
Biometric sample data as WM host ("sample watermarking") vs. embedding of biometric templates into arbitrary or biometric cover data ("template embedding"). Here: BOTH !!

- Steganographic approach: The biometric data to be transmitted is hidden into a carrier image where the aim is to conceal the transmission of the embedded biometric data.
- Sample-replay prevention: When acquiring sample data, these are robustly watermarked, such that sniffed data of this type cannot be used to fool the sensor pretending these to be real data.
- Multibiometric approach: A host-image, e.g. fingerprint, taken by a sensor at the authentication point is used in conjunction with another biometric, e.g. iris, from the same user (eventually stored on a smart-card which has to be submitted by the holder at the access control site).
- Sensor and Sample Authentication approach: A WM is used to ensure the integrity of transmitted biomentric sample data and the entire authentication chain.

Two-Factor Biometric Recognition with Semi-fragile Template Embedding

We focus on a two-factor authentication scheme based on biometrics and a token, i.e. a smart-card. When a user is enrolled into the system, sample data are acquired, corresponding (enrollment) template data is extracted and stored in two different ways:

1. In the centralized biometric database required for the actual recognition process and
2. On the smart-card as submitted by the user for initiating the verification.



Verification Process

1. From the acquired sample data, a template is extracted and compared to the template on the smart-card (without contact to the central database). Only if there is sufficient correspondence, the following stages are conducted subsequently.
2. The smart-card embeds its enrollment template 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 is extracted, and is compared to the template extracted from the sample (denoted simply as "template" in the following) to check the integrity of the transmitted sample data.
5. Finally, in case the integrity of the data has been proven, the template watermark and the template are used in the matching process, granting access if the similarity to the enrollemnt template in the database is high enough.

Comparison to Related Techniques

- Robust WM embedding: sample data integrity is ensured in addition to sole transportation (see below).

- Common semi-fragile WM: we do not need to know the WM at the receiving side and the embedded data can be immediately used for improving matching.

- Digital signatures: a certain amount of robustness is given as well as the position of eventual tampering locations; there is no additional data like the signature itself and a digital signature cannot support both, integrity and the two-factor approach.

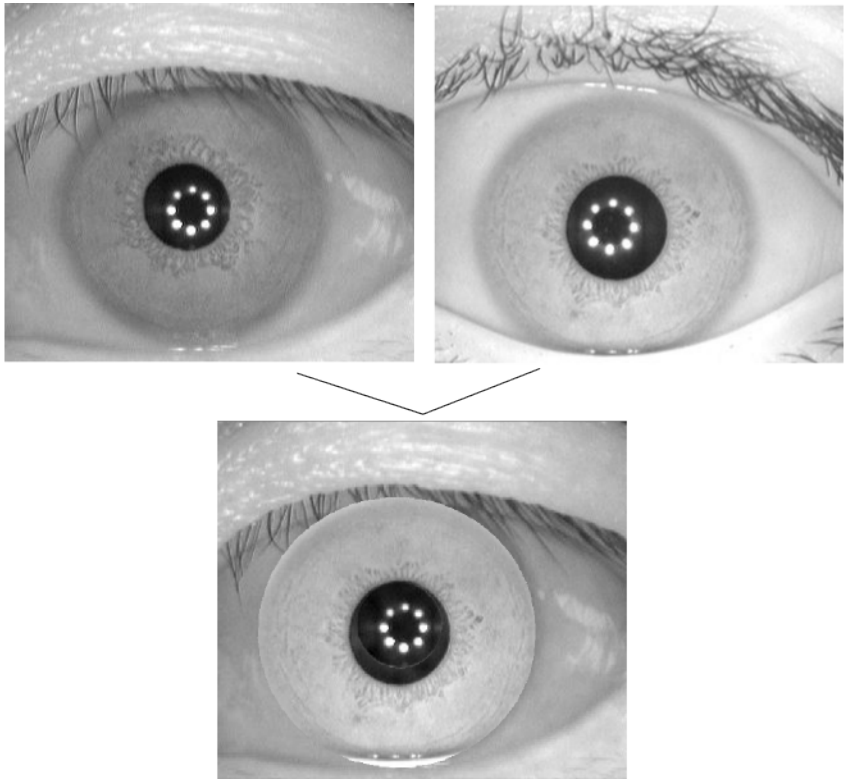
Attack against the ROBUST Watermark Case

The System: We focus on a robust watermarking approach enabling two-factor authentication where data stored on a smart-card is embedded into iris sample data.

Attack assumptions: We suppose the attacker can utilise a stolen smart-card to fool the system. Additionally, he is in possession of sniffed sample iris data of the person owning the smart-card (the legitimate user) which could have been acquired with a telephoto lens or cropped from his high-resolution personal Facebook image for example.

The Attack: The attacker uses the biometric system pretending to be a legitimate user: the smart-card is inserted, an iris sample is acquired, and finally, the data stored on the smart-card is embedded into the iris sample. Now the attacker intercepts the transmission of the data to the matching module. He modifies the iris image such that the attackers' sample data matches that of the sniffed sample data of the legitimate user while not destroying the embedded WM information (it seems reasonable to assume this capability since use of robust WM suggests some public channel).

Visual Attack Example



The iris texture of the left image (attackers' sample) is replaced by the iris texture of the right image (legitimate users' sniffed sample data), thus resulting in a new iris image as shown in the figure (still watermarked with the legitimate users' template).

Experimental Settings

- Iris Recognition Software: Libor Masek's Matlab implementation of a 1-D version of the Daugman iris recognition algorithm.

- Iris Databases:

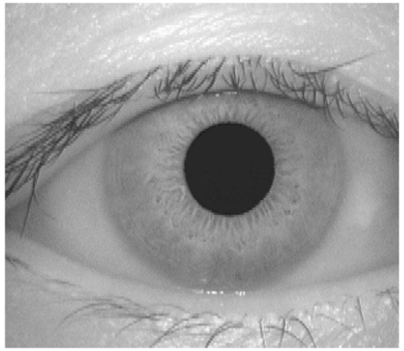
  CASIAv3 Interval database out of which 500 images have been used in the experiments.

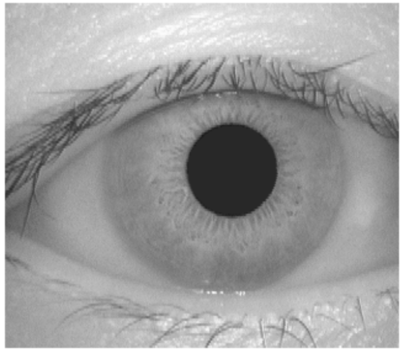
  UBIRIS database out of which 318 images have been used in the experiments.

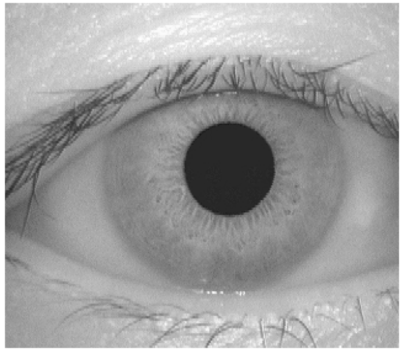
  MMU database consists of 450 images which all have been used in the experiments.

- Watermarking scheme: Fragile Watermarking scheme by Yeung et al. with capacity of 89600, 76800, and 30000 bits for CASIAv3, MMU, and UBIRIS, respectively.

Visual Results: Tamper Detection



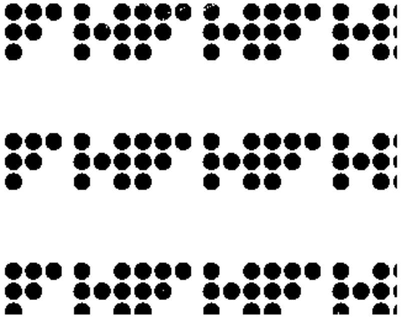


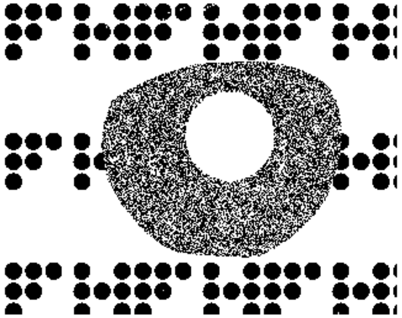


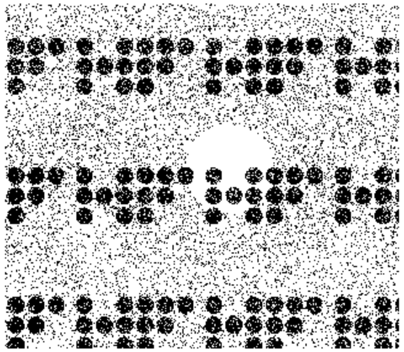
Original

Replaced iris

JPEG compression







Extracted Watermarks

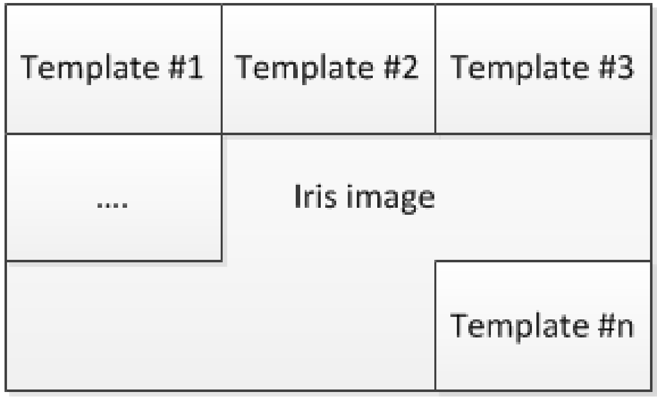
Robustness or the Original Yeung Scheme: Watermarking Bit-Error-Rate (BER)

Attack	CASIAv3	MMU	UBIRIS
Mean filtering	0.50	0.50	0.50
Gaussian Noise $N = 0.0005$	$4.6 \cdot 10^{-5}$	$5.6 \cdot 10^{-5}$	$6.1 \cdot 10^{-5}$
Gaussian Noise $N = 0.001$	0.03	0.03	0.03
JPEG Q100	0.05	0.06	0.05
JPEG Q95	0.43	0.45	0.45
JPEG Q75	0.49	0.50	0.50

$\longrightarrow$ Some limited amount of robustness against 100% JPEG and noise only.

Redundant Embedding

The smaller size of biometric templates can be exploited to embed the template in redundant manner: the 9600 bits templates can be embedded 9, 8, and 3 times into images from the CASIAv3, MMU, and UBIRIS databases, respectively.



Template #1

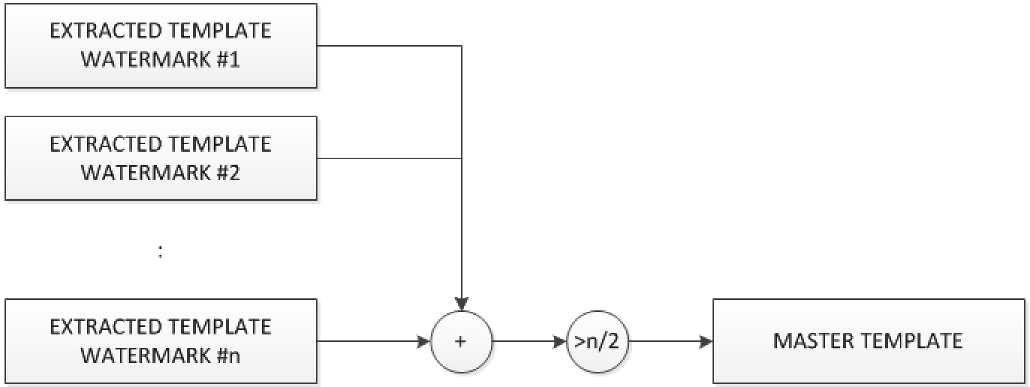
Template #2

Template #3

....

Iris image

Template #n



EXTRACTED TEMPLATE WATERMARK #1

EXTRACTED TEMPLATE WATERMARK #2

:

EXTRACTED TEMPLATE WATERMARK #n

+

>n/2

MASTER TEMPLATE

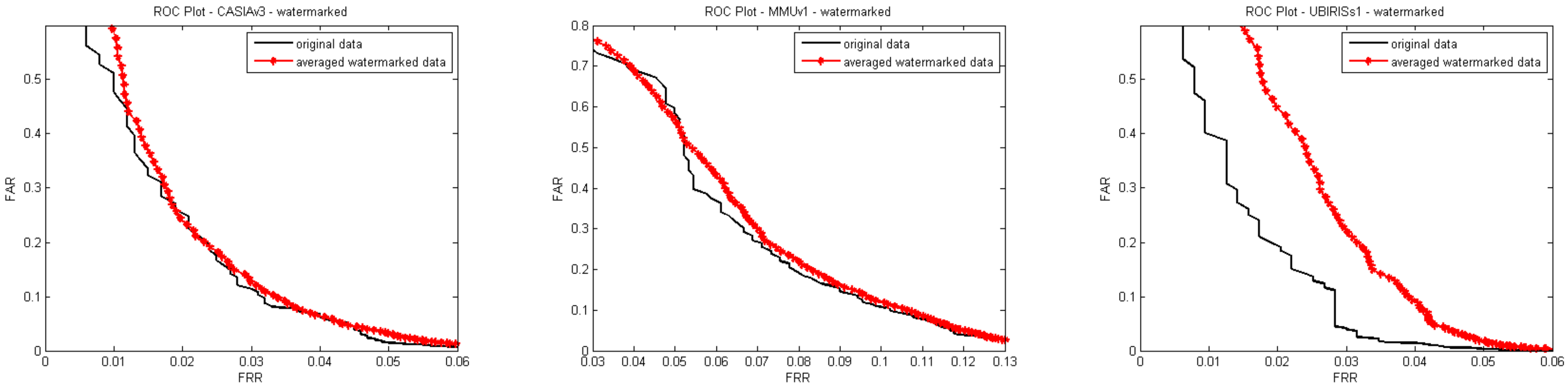
Robustness for Redundant Embedding: Template BER

Attack	CASIAv3	MMU	UBIRIS
Mean filtering	0.50	0.50	0.50
Gaussian Noise $N = 0.0005$	0	0	0
Gaussian Noise $N = 0.001$	0	0	0.003
JPEG Q100	0	0	0.01
JPEG Q99	0	0.01	0.05
JPEG Q98	0.08	0.14	0.22
JPEG Q95	0.35	0.40	0.43

$\longrightarrow$ we notice increasing robustness for an increasing amount of redundancy (CASIAv3 has maximal redundancy, i.e. 9 times).

WM Impact on Recognition Performance

Original ROC performance is compared against recognition using watermarked data (the average of ten embedded WM is shown).



CASIAv3, MMU, UBIRIS

⟶ while for the CASIAv3 and MMU there is hardly a noticeable impact, we notice significant result degradation in the case of the UBIRIS dataset. This is due to the already low quality of this dataset, in case of additional degradation results get worse quickly.

WM Impact on Recognition Performance under Attacks

Beside the EER, we show FRR (for $FAR = 10^{-3}$) and FAR (for $FRR = 5 \cdot 10^{-3}$).

		ERR	FRR	FAR
CASIAv3				
no attack	original	0.045	0.091	0.650
	template watermark	0.048	0.081	0.742
mean filter	original	0.035	0.061	0.644
	template watermark	0.044	0.063	0.669
JPEG Q98	original	0.037	0.074	0.626
	template watermark	0.049	0.086	0.617
UBIRIS				
no attack	original	0.032	0.062	0.764
	template watermark	0.046	0.071	0.865
Gaussian Noise $N = 0.001$	original	0.038	0.068	0.871
	template watermark	0.049	0.073	0.868
JPEG Q95	original	0.036	0.066	0.838
	template watermark	0.045	0.070	0.975

$\longrightarrow$ in any case, we notice a slight result degradation for the variant with embedded WMs.

Robust Integrity Verification

We measure BER between the template WM and a database template that has been generated by majority voting among 5 different templates. A typical decision threshold for the iris recognition system in use is at a BER ranging in $[0.3, 0.35]$.

Attack	CASIAv3	MMU	UBIRIS
No attack	0.21	0.23	0.19
Mean filtering	0.49	0.50	0.50
Gaussian Noise $N = 0.0005$	0.21	0.23	0.19
Gaussian Noise $N = 0.001$	0.21	0.23	0.19
JPEG Q100	0.21	0.23	0.19
JPEG Q99	0.21	0.24	0.22
JPEG Q98	0.25	0.30	0.32
JPEG Q95	0.41	0.45	0.45

$\longrightarrow$ we realize that integrity verification in our technique is indeed robust against moderate JPEG compression and noise. However, mean filtering and JPEG compression at quality 95% destroys the template WM and indicates modification.

Conclusion

- There are many different proposals how to use WM in the context of biometrics. In many schemes, the used WM technology does not fit well the requirements of the biometric system.

- When using WM as a sole means to enable a two-factor authentication scheme, semi-fragile or fragile WM techniques can prevent cropping attacks and can provide (semi-fragile) integrity verification. The distribution of incorrect bits can be further used to differentiate between malicious attacks (where an accumulation of incorrect bits can be observed in certain regions) and significant global distortions like compression.

- Contrasting to claims in literature, recognition performance of the templates extracted from watermarked sample data suffers from degradation to some minor extent, even for the considered fragile embedding scheme. However, this can more than compensated by the additional template watermark which should be involved in matching as well.

Thank you for your attention !

Questions ?