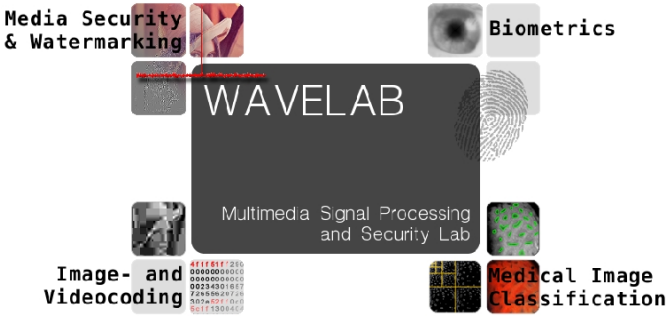
Watermarking as a Means to Enhance Biometric Systems: A Critical Survey

Andreas Uhl

Department of Computer Sciences
University of Salzburg, Austria

uhl@cosy.sbg.ac.at

http://www.wavelab.at/



UNIVERSITÄT
SALZBURG

UNIVERSITÄT SALZBURG

Outline

- Introduction & Motivation
- Watermarking in Biometrics
- Application Scenarios for WM in Biometrics
- Discussion of Application Scenarios
- Open Issues & Conclusion

Introduction

With the increasing popularity of biometric recognition applications, several security breaches have been discovered.

Strategies to cope with some of those issues are template protection schemes (based on cancelable biometrics or biometric cryptosystems), cryptographic techniques to provide matching in encrypted domains (like homomorphic encryption schemes), liveness detection strategies, or other means of protecting the integrity of the entire authentication mechanism.

Watermarking (WM) has been suggested as a means to resolve some of these problems as well and can potentially add additional functionalities to biometric systems.

In this context, the overall impression arises that two "buzzwords" (i.e. watermarking & biometrics) have been combined without carefully analysing what this combination should actually achieve, why, and how this could be done in a sound manner.

Watermarking and Biometrics

Biometric sample data as WM host ("sample watermarking") vs. embedding of biometric templates into cover data ("template embedding").

Biometric watermarking: The aim of watermarking is not to improve any biometric system, but to employ biometric templates as "message" to be embedded in classical robust watermarking applications like copyright protection in order to enable biometric recognition after the extraction of the WM. The most famous example is the "secure digital camera" where an iris template of the photographer is embedded into digital images. This has been suggested for various types of media (e.g. including 3D mesh data).

Impact on recognition performance: in case a WM is embedded into sample data to be subsequently used for recognition, the potential impact on recognition performance is an issue (e.g. impact has been reported for iris, speech, and fingerprint recognition). A possible strategy to cope with this is to protect the sample areas during the WM process, which are of importance for the subsequent recognition process (like sparing out fingerprint feature regions close to minutiae for fingerprints). Another approach is the application of reversible WM schemes.

Watermarking Application Scenarios in Biometrics

- Covert (Template) Communication: 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.
- 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.
- Two-Factor authentication: Authentication data of a second type is embedded into sample data using WM technology (eventually stored on a smart-card, as an alternative classical PWD information can be used).
- 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.
- Sensor and Sample Authentication approach: A WM is used to ensure the integrity of transmitted biomentric sample data and the entire authentication chain.

Watermark Properties for Biometric Applications

In realistic biometric application scenarios there is usually no unmarked original image available to detect the WM in the marked image (contrasting to many DRM scenarios). Therefore, only **blind** WM techniques are applicable in most cases.

A significant amount of techniques proposed in literature employ **robust** blind WM embedding techniques (i.e. robust against unintentional image manipulations). It is crucial to determine if robustness is indeed required in the sense of DRM scenarios since robust techniques have been found to impact recognition performance more as compared to most (semi-)fragile schemes.

The application scenarios considered differ significantly in terms of the amount of data to be embedded – while some employ zero-bit WM techniques (to be able to reliably detect WM presence), some only aim at embedding a sensor ID to authenticate an admissible biometric sensor, others even aim at embedding biometric template data (e.g. 2048 bits in case of the iris code). Therefore, the **capacity** of the employed WM techniques is important.

Covert (Template) Communication

The aim of watermarking in this approach is to transmit biometric data (template or even sample data) hidden into arbitrary carrier / host data (an attacker should be unaware of the concealed (real) data transfer). Therefore, this is a typical *steganographic* application scenario, which is based on template (or even sample data) embedding.

Attack An attacker aims at *detecting* the WM in order to be able to intercept the template data transfer.

WM properties and content As a consequence, the WM has to be capable of carrying the template / sample data (capacity requirement) and has to be undetectable. In the passive warden scenario, robustness of the WM is not an issue, however, robustness contradicts the requirement of a non-detectable WM. Blind extraction is required as it is a must for all steganographic application scenarios.

Crypto alternative There is no cryptographic technique capable of replacing a steganographic approach.

Covert (Template) Communication: Questions

- In most biometric systems, a biometric sensor is typically expected to transmit biometric authentication data over a dedicated channel to the feature extraction / matching module. So the gain of steganography is not clear. Only in a distributed biometric system where authentication data is transmitted over networks where also other type of data is communicated the approach makes sense.

- Almost all proposals in literature suggest to use robust WM for embedding, which which actually destroys the steganographic non-detectability property. The remaining value of the proposed robust schemes is in communicating embedded templates in a way that the are not **perceived** by a human observer. The result is a weak concealment of template content avoiding encryption. In this manner, neither the steganographic nor the confidentiality aim can be met. When applying robust embedding as being proposed, embedded templates resist non-malicious cover data manipulations (which is an advantage over steganographic schemes in the case of an active warden). Thus, the application context has to determine if robust WM or steganographic embedding serves the actual aim of the WM embedding.

Multibiometric Recognition

The aim of watermarking in this scenario is to embed biometric data (sample data - large data volume vs. template data - extraction required) into a biometric sample in order to facilitate the employment of a multibiometric technique. The aim is an increased recognition performance.

Attack The resulting system is vulnerable in principle against all types of attacks endangering classical unimodal systems systems. In particular, an attacker needs to $embed$ sniffed biometric data of one modality into sniffed sample data of a second modality.

WM properties and content As a consequence, the WM has to be capable of carrying either template or sample data (capacity requirement) and extraction has to be blind. It is of advantage if the WM resists unintentional image manipulations like compression or noise insertion, but robustness is not a required property here. In order to prevent an attacker to embed a stolen template, the embedding algorithm could be dependent on a key.

Crypto alternative The benefit of embedding additional authentication data with WMs over classical cryptographic schemes is that this may be done in a way where "allowed" manipulations can be conducted on the data.

Two factor Authentication

The aim of watermarking in this scenario is to embed the data corresponding to a second authentication factor into a biometric sample. The aim is to increase security by introducing an additional but different authentication scheme.

Attack The attacker can utilise a stolen smart-card (or sniffed password) and additional sniffed sample data of the attackers' target subject to fool the system. He uses the biometric system pretending to be a legitimate user, but after WM embedding (e.g. of the data stored on the card), the attackers' sample data is $tampered$ to match that of the sniffed sample data while not destroying the WM.

WM properties and content The WM has to be capable of carrying the additional authentication data (capacity requirement: passphrase, ID, or template data) and extraction has to be blind. In order to resist against a manipulation of the attackers' sample acquired by the sensor the WM scheme employed must not be robust. Therefore, only semi-fragile or fragile WMs fit all requirements.

Crypto alternative The situation is perfectly identical to the multibiometric scenario and shares all corresponding problems discussed before.

Multibiometrics & Two factor Authentication: Questions

- In both scenarios, WM is used as a means of transportation of two different data sets in a tightly coupled manner. Application of encryption to a concatenation of both data pieces results in slightly more data to be transmitted, but unauthorised embedding into or manipulation of transmitted biometric data is prevented by this technique. Furthermore, this approach definitely has no impact on recognition performance as opposed to WM embedding. So it is highly questionable, if the idea of using WM for transport is a good one.

- Most schemes propose to use robust WMs for embedding and therefore do not provide the capacity for sample embedding (for many modalities, not even for template data). It seems that for this application case, it would therefore be better to abandon the idea of providing robustness but to use fragile or steganographic embedding techniques (eventually protected by error correction coding to provide some limited resistance against channel errors or lightweight signal processing). Additionally, the recognition degradation and vulnerability against tampering can be handled in this manner.

Sample Replay Prevention

During data acquisition, the sensor (i.e. camera) embeds a watermark into the acquired sample image before transmitting it to the feature extraction module. As soon as an intruder presents the sniffed biometric sample data to the sensor, it can detect the watermark, will deduce non-liveness and will refuse to process the data further.

Attack An attacker aims at *removing* the WM in order to be able to use sniffed data for replay attacks or as fake traits.

WM properties and content As a consequence, the WM has to be robust. It has to be detectable in the image as long as the image can be used in the recognition process. The extracted mark needs to carry at least the information "yes, I have been acquired by a sensor" (so eventually zero-bit WM could be used), but could also carry actual sensor IDs.

Crypto alternative Encrypting the data after acquisition for transmission provides similar functionality, however, the data needs to be decrypted for feature extraction and matching, which is a severe disadvantage. In any case, the WM may serve as additional "second line of defence" as it is suggested in the DRM context as well.

Sample Replay Prevention: Questions

- The approach only works if sniffed data is used in the biometric system it have been sniffed from. Sample data acquired from a different database or acquired even with some other sensors (e.g. consumer camera) will not be detected. Generic liveness detection techniques also target the attempt of using sniffed image data to fool the sensor and are much more generic.

- Since embedding has to be done in a robust manner, impact on recognition performance has to be expected.

- Questions of WM security are equally important as for robust WM in the DRM area.

Sensor and Sample Authentication

The aim of watermarking in this scenario is to ensure the integrity 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.

Attack An attacker aims at *inserting* the WM in order to mimic correctly acquired sensor data.

WM properties and content In contrast to the previous scenario, the WM needs to be unique in the sense that it has to uniquely identify the sensor and carry a unique transaction number or timestamp. Resistance against a WM insertion attack can be achieved by sensor-key dependent embedding. Since the watermarking scheme has to be able to detect image manipulations, (semi-)fragile techniques are the method of choice.

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.

Sensor and Sample Authentication: Questions

- The discussion if WM is a sound alternative to more classical authentication schemes for media data is not at all specific to the biometric scenario, all respective arguments of the general discussion are valid here. For example, digital signatures represent separate data which has to be taken care of separately, and are usually not capable of providing any robustness against channel errors and unintentional signal processing. Additionally, WM eventually provide information about the location where image data tampering has occurred.

- Potential impact on recognition performance has to be considered (but is expected to be of minor importance due to the employed Wm techniques).

- Information about the location where sample data tampering has occurred can be important for the assessment if the tampering has to be considered significant.

Conclusion

- The WM schemes as suggested to be used in the context of biometric systems often exhibit somewhat adhoc properties and specific requirements are not analysed in detail. In many cases, the actual WM method proposed does not lead to the desired effect or at least not in an optimal manner.

- For most scenarios considered, WMs are not the only means to achieve the desired goals (and for some scenarios, WM are definitely not the best means to do so).

- More thorough investigations are required in this field to (a) identify sensible application scenarios for watermarking in biometrics (e.g. exploiting similarity of biometric templates and robust image descriptors as used in semi-fragile WM) and to (b) select and/or design appropriate WM schemes to support the desired functionalities better (e.g. reversible schemes with appropriate capacity).

Thank you for your attention !

Questions ?