CHAPTER 1

INTRODUCTION

This chapter includes the following subtopics, namely: (1) The Rationale: (2) Theoretical Framework; (3) Conceptual Framework and Paradigm; (4) Statement of the problem; (5) Hypothesis; (6) Objective; (7) Scope and Delimitation; and (8) Significance of the study.

1.1 Rationale

Authentication system was mostly used in security applications and user identification system. It was used to be able to access or use certain services. The example of services using authentication system which was used to verify identity can be found in gadget lock (computer or mobile device log-in) [12], door accesses in the hotel, apartment, as well as certain agencies, on-line payment, and others. In the early development, the authentication systems still used mechanisms such as username and password for authentication. But at the present time, the authentication systems are already widely designed and manufactured based on biometrics. The goal of these later development was to determine or verify the identity of users using physiological and behavioural biometrics including faces, fingerprints, irises, retinas, voices and others [8].

Recently, authentication applications based on face identification are becoming popular and are increasingly developed since biometric data of the faces (photos, videos) can be easily taken with the available devices like cameras, not like other biometric sensor technologies. One biometric data of faces can also be used in many different environments. Therefore, face authentication has been widely used in identification system and access management such as bankcard identification, access control, security monitoring, and surveillance system.

However, this type of authentication was very vulnerable to spoofing attacks. Spoofing attack was a type of attack wherein an attacker presents a fake biometric data to the acquisition sensor with the goal of impersonating him/herself as a legitimate user to get illegitimate access and advantages. Spoofing attack (or copy attack) was still a fatal threat for biometric authentication systems because the faces are visible, the voices are recordable, and fingerprints are left everywhere the persons go [13]. In face authentication system, spoofing attacks mainly consist of print attack and replay attack. Print attack uses printed photographs of a subject or original user, while replay attack presents a video of a live subject by playing video or by displaying digital photo in front of the camera [1]. These

actions can be easily produced since a lot of multimedia contents, i.e photographs and videos, are openly available on the internet because of the increasing popularity of social network sites (Facebook, YouTube, Instagram and others). This indicates that there are many facial image data which can be used for falsification of identity with a face spoof attack so that attackers can use the data of one original user to access a service, visual surveillance, compute application security or for business transactions.

With the increasing popularity of face authentication on one hand and so many facial image data available in social network sites on the other hand, the need to incorporate face spoofing detection tool into face authentication becomes significant.

1.2 Theoretical Framework

Spoofing attack in the face authentication systems usually uses a photo, 3D mask or video of a valid user. Print photo attacks (Figure 1.1 a) require the use of high quality 2D printers and 3D facial mask (Figure 1.1 b) require high resolution fabrication capturing the 3D shapes of the valid users face [12]. Video replay attacks with displaying a video from video or photo (Figure 1.1 c) are easier to launch than either printed photo attack or 3D mask attack and it was difficult to detect video replay attack since it was a re-imaging of the original live face and high quality of this attack was almost the same as live faces. The video replay attack has more challenges than print photo attack (static image). This was a type of attack which was most easily fool the camera, particularly if the attacker uses the advanced recorder device to spoofing. It was a big threat to face authentication system, because this attack was very difficult to distinguish visually. Video replay attack has more physiological clues than photos, such as eye blinks, facial expressions, head movements and shoulder movements. So it was necessary to develop spoofing detection module mainly for video replay attacks.

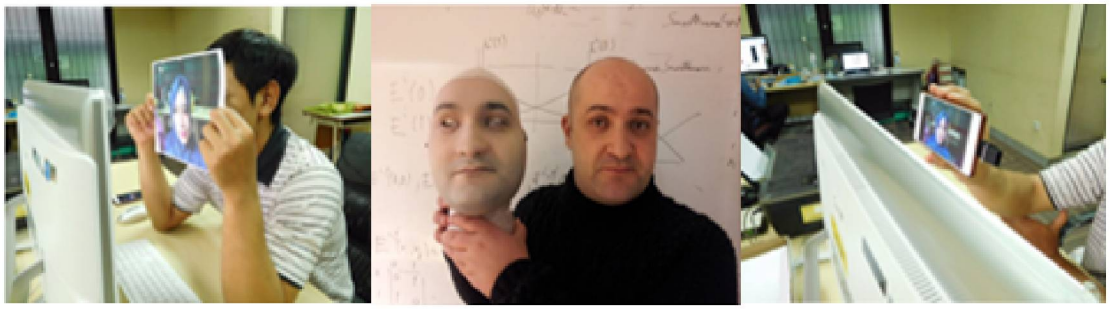

Figure 1.1: a. Print photo attack; b. 3D mask attack; c. Replay attack (Photo/Video)

Since the spoofing attack was a major issue in the authentication system, spoofing detection module should be a part of the authentication system which should be applied before the identification or verification module. This was illustrated on the global scheme of the authentication process in Figure 1.2.

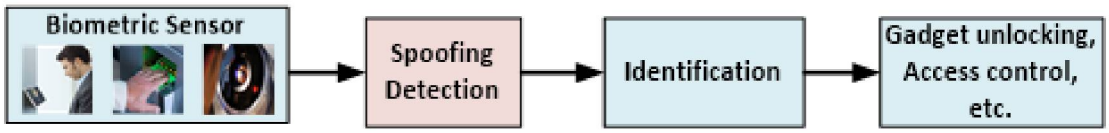

Figure 1.2: A spoofing detection on the biometric authentication system

In authentication system, when a client uses his/her biometric data (face, fingerprint, iris, etc.) as input to access certain services, the system must verify authenticity of the data in front of the sensor. This system will detect whether the input was spoofing or live biometric data, and if it was accepted as the live user, then the system proceeds to the identification process by matching the input with the existing user's database. When input data was identified, then the user can use the access service he/she needs.

The most commonly studies on face spoof detection use facial motion approach such as eye blink, mouth movement, facial expressions, and head movements [1, 2, 9–11, 17, 20]. The previous research in [9] has developed the spoof detection using fusion of motion and texture analysis. Now so many accesses in security system involves the motion analysis for face authentication [1].

1.3 Conceptual Framework and Paradigm

This study focused only on the development of spoofing detection module for video replay attack on the face authentication system as was shown in Figure 1.3.

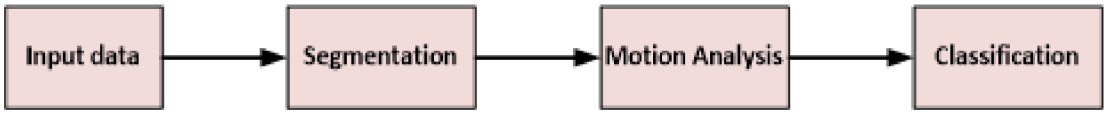

Figure 1.3: A proposed scheme of video replay attack detection on the face authentication system

The steps of the scheme are: (a) inputting data consisting of real videos and spoof videos; (b) separating process between the background and the object (head and shoulders) areas by model-based segmentation; (c) motion analysis, conducted not only on the face area but

also through the whole frame area including the background; (d) classification to determine real or spoof video input data.

The model-based segmentation requires a user be within the model to ease separating process of the object from the background. The definition of the object here was the upper of the user's body. In the real application, the proposed method useful for face authentication that utilizes a webcam on the computer or the laptop such as for the access control (lock and unlocks) computer. The webcam that was embedded in the computer could allow a user to be in front of the camera (the center of view camera) with a certain distance so that the upper of the user's body was caught, so it would preserve the accuracy of the system.

1.4 Statement of the Problem

As aforementioned, Komulainen et al [9] has developed the spoof detection using fusion of motion and texture analysis. In their work, a square area around the middle of camera view has been determined for the face area. So, users/clients must adjust his/her face to be within the square area. This area was used to segment the face area from other, and then the motion information from the whole area was analyzed while the textual features was extracted only from the square area where the face was supposed to exist. The consequence of this scheme was that the motion of the body of the object outside the square area such as the shoulder and hair will not be handled, while the textual features can also appear in the complex (varied, miscellaneous) background condition.

Figure 1.4: (a) an image with homogeneous background; (b) an image with complex background

The complex background as shown in Figure 1.4 (b) has a composition of unrelated element or part, other artifacts (such as: a window frame, and other objects that have textures) other than a single color wall. So, if the background condition was complex, it was very difficult to distinguish between real and spoof using texture analysis.

1.5 Objective

The objective of this study was to build a tool that can detect video replay attack in indoor environments with static background and lighting conditions from indoor and outdoor with improved the error rate of the previous research [9].

1.6 Hypothesis

In the spoof/attack video, motion patterns appear in the whole frame or no motion at all, while in real video, motion patterns only appear in the object area (head and shoulder) and there was nearly no motion in background area. Therefore, by using motion analysis in the whole frame of the video (background, head and shoulder), the difference between spoof and real video can be identified.
Different from the previous research [9], this study uses a model for segmentation process to distinguish between the background and the object/user area, so that all parts of the object/user area can cover hair and shoulder, instead of the only predetermined face area as proposed by the previous study. The model-based segmentation requires a user to be within the model to ease separating process of the object from the background.

1.7 Scope and Delimitation

The experiment using IDIAP Replay Attack dataset from 50 subjects consisted of 200 real video clips and 1000 spoof video clips. While for additional experiment using personal synthesized dataset also from 50 subjects, consisted of 100 real video clips and 400 spoof video clips with the lighting condition from indoor and outdoor. In this study, the tool could only be used in indoor with static (no movement) background, and with a moving object/user. The object/user was always in the center position of the sensor camera coverage. The motion to be analyzed in this study had no specific meaning such as eye blink, and mouth movement, but the motion here was defined as a changes of spatial features between frames.

1.8 Significance of the Study

In this study, the tool may increase the security of the authentication system because an attack was expected to be detected before the identification process, so as to facilitate the authentication system and reduce the potential for forgery. The error rate can be improved by taking different approach from previous research [9] which was apply a model for segmentation process to detect the object/user area. The model covers not only the face (as in previous research [9]), but also the entire head and shoulders, and with only using motion analysis (instead of using motion and texture as has been adopted in previous study), the computational cost can be reduced while at least the accuracy of the system can still be preserved.