SOCIETY FOR SCIENCE AND EDUCATION
UNITED KINGDOM

TMLAI TRANSACTIONS ON
MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Volume 5 No. 2
ISSN 2054-7390

Face Spoofing and Counter-Spoofing: A Survey of State-of-the-art Algorithms

[1]Rinku Datta Rakshit and [2]Dakshina Ranjan Kisku
[1]Department of Information Technology, Asansol Engineering College
Asansol, West Bengal, India
[2] Department of Computer Science and Engineering, National Institute of Technology Durgapur
Durgapur, West Bengal, India
rakshit_rinku@rediffmail.com; drkisku@cse.nitdgp.ac.in

ABSTRACT

In the current scenario of biometric-based identity verification, a face is still being proved to be an essential physiological evidence for successful person identification without letting know the target. Nevertheless, repeated attacks of intruders can cause the face recognition system insecure because of easy availability of face images or pictures of a person in social networks or in other networked resources. Spoofing facial identity in a biometric system is not a difficult task for intruders. When an intruder presents a photograph or a video containing a face of a person in front of a networked camera which is integrated with a face biometric system, spoofing is referred to as presentation attack. Without anti-spoofing mechanisms, biometric systems are at high risk in case of susceptible attacks. Thus, detecting face spoof in a face biometric system is a challenging research field. The aim of this paper is to summarize some of the most popular face spoof detection techniques which are proved to be beneficial for the researchers to make it an indispensable aspect in biometrics.

Keywords: Biometrics, Face biometric, Face spoofing, Face anti-spoofing, Printed photo attack, Replay video attack, 3D mask attack, Plastic surgery attack, Texture analysis, Motion analysis, Liveness detection, Image quality analysis.

1 Introduction

In the area of person identification, biometric-enabled applications are increasing day by day because it is found much more secure than any other security means such as password, token, PIN or key, as each person have their unique characteristics to be identified themselves correctly. The face is an important physiological evidence which is very popular in biometrics because of its non-intrusiveness and simplicity in use. Despite several advantages of a biometric system, it may get worse day by day because of spoofing face evidence done by some unauthorized attackers and intruders. Attackers find various ways to spoof the face data from a biometric system. Easy availability of face images or face videos of a person in social networks may take in intruders to take advantage of face spoofing in a biometric system. They present the face images or face videos at the face of the camera and try to spoof the system. Due to advancements in medical science, plastic surgery of a face could be a possible threat to a biometric system. The 3D face mask as well helps the attackers to spoof the face data in this area. Thus, consolidation of a powerful face

DOI: 10.14738/tmlai.52.3130
Publication Date: 09th May, 2017
URL: http://dx.doi.org/10.14738/tmlai.52.3130

spoof detection technique with existing biometric system can be achieved and this requires to be addressed urgently.

Currently, biometric researchers are working towards this specialized area to integrate such novel ideas together to make robust anti-spoofing systems. TABULA RASA (Trusted Biometrics under Spoofing Attacks) [49] is a European Commission funded project working under the Seventh Framework Program. This project addresses some of the issues of spoofing attacks to trusted biometric systems.

Different techniques for spoof detection [1-8], [10], [14], [50] are introduced by different researchers in this area. The authors in [1] proposed a liveness detection technique using motion estimation which is able to detect 2D spoof attacks. Motion estimation is done using optical flow. Experiments on a private database showed that 6% false-alarm against about 14% false-acceptance. An optical flow based method to capture and track the delicate movements of different facial parts is presented in [2] assuming that facial parts in real faces can move differently than in photographs. The method suggested in [14] differentiates the 2D images of face prints and 3D live faces by considering the Lambertian reflectance. Latent reflectance features are extracted in this method using a variational retinex-based method and Difference-of-Gaussians (*DoG*) based approach. The latent reflectance features are then applied for classification. Geometric invariants are used to detect replay attacks, proposed in [4] using two video databases. A score level fusion approach using LBP, the histogram of oriented gradients and Gabor wavelets computed from the local blocks of a face image is given in [5]. The histogram computed from all the blocks are concatenated later for three different feature descriptors and further, it generates three feature vectors. Kernel approximations of each of the three feature vectors are computed, and a linear support vector machine (SVM) is used for classification. In the end, the match scores of all three SVMs are fused to produce the final results. The authors report in [6] that HTER on the Print Attack dataset is 0%. LBP from three orthogonal planes (LBP-TOP) is applied to detect spoofing in the Replay Attack database. LBP-TOP exploits the temporal information by computing LBP histograms in the XT and YT planes along with spatial data in the XY plane. This approach achieved best HTER of 7.6%on the Replay Attack dataset. However, LBP-TOP is found computationally expensive and may not scale to real-time applications. In a recent work discussed in [7] is an extended work of the technique proposed in [3] to an image-based spoofing detection based on the fact that the brightness of the LCD screen affects the recaptured image, which makes the image edges more susceptible to a blurring effect. This approach brings in one new intermediate step. Adaptive histogram equalization is done to the image before extracting latent reflectance features. The reported results in the publicly available NUAA database and Yale face database show that the proposed extension reduced the classification error in more than 50% for high-quality printing spoofs in the NUAA database and 65% for images recaptured from an LCD monitor for the Yale face database. Faces are described by a combination of several feature descriptors such as shape, color, and texture in the work explained in [8]. This work provides improved results, but the combination of these descriptors generated high dimensional feature spaces that may not be suitable for standard classification methods. An eye blinking based work is proposed in [50]. It is believed that a person blinks about once every 2 to 4 seconds and it is averred in the work proposed in [39]. This study uses an undirected conditional random field framework to represent eye blinking from Hidden Markov Models that relax the independence assumption of generative modeling, with the advantage that the method allows to relax the assumption of conditional independence of the observed data. In [10], the contributors are extended the study introduced in [50]. This work introduced a scene context matching in stationary face recognition systems. For this, the authors analyze inside face cues such as eye blinking and outside face cues, since the background scene is known by the recognition system. The authors reported that their method works well to photo-based, video-based and 3D-based spoofing techniques. To detect photo-based spoofing and 3D models, spontaneous eye blinking inside-face clues can be employed and to detect video-based spoofing, outside-face clues of scene context are used. Nevertheless, since a video or image background can be easily altered, their method may fail and a 3D model can incorporate the natural process of blinking similar to a real eye blinking. A private dataset, in which they obtained almost perfect results, is created but not released to the public. Though a bit of face spoof detection techniques has been aimed, still their generalization skill has not been adequately addressed.

In spite of several advancements, researchers are even facing some problems due to the lack of publicly available face spoofing databases which hold different types of spoofing attack examples. However, NUAA [14] and Print-Attack [22] both are publicly available face spoofing databases which contain only printed attacks. CASIA-FASD [9] database contains printed photo, printed photo with perforates eyes regions and replay video attacks. These databases are not found very useful because, lack of precise database protocol which describes train, development and test set clearly. Therefore, very limited number of studies on face-anti spoofing methods are available.

The paper is organized as follows. Section 2 describes some useful terminologies related to biometrics and face spoof detection techniques. Research challenges are introduced in Section 3. Section 4 explains the taxonomy of face spoof detection techniques where categorizations of face spoof are given. Existing face spoof detection algorithms, and their pros and cons are discussed in Section 5. Section 6 reports the competitions held and their results on face spoof detection. Section 7 presents different databases which are available on face spoof detection. Comparison of different face spoof detection algorithms is presented in Section 8. Concluding remarks are made in the last section.

2 Some Useful Terminologies

2.1 Biometrics Traits

Biometrics refers to machine vision technology designed for distinguishing a person based on one or more physical or behavioral features and it is widely used in access control and security systems. A biometric system can use face, voice, fingerprint, gait, retina and iris of a person for identity verification. Biometric systems can be categorized into two groups, such as unimodal and multimodal biometric systems.

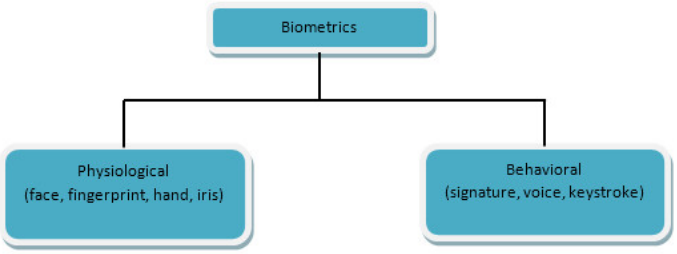
Biometrics
Physiological
(face, fingerprint, hand, iris)
Behavioral
(signature, voice, keystroke)

Figure 1. Different types of biometrics.

A unimodal system uses single evidence (physiological or behavioral) of a person for identification. It is easy to delude. A multimodal system uses more than one biometric evidences of a person for establishing the individuality of that individual. It is more secure than the unimodal system. Two types of biometric evidence are shown in Figure 1.

Amongst biometric modalities, face biometric systems are deployed in large number for various security applications such as access control, surveillance, and criminal identification. Face biometric is natural, intuitive, easy to use and less human – invasive.

2.2 Face Spoofing Sources

Face spoofing is a grave threat to identity stealing in face recognition. Face spoofing is a means by which an illegitimate person tries to spoof a face biometric system by presenting a face image, video or a face mask of a legitimate user at the face of a camera to gain access to the secured resources. One can also use make-up or plastic surgery as a means of face spoofing, but the most common sources of face spoofing are photographs and videos. As the popularity of using social network websites (Facebook, Flicker, YouTube, LinkedIn and others) are increasing among a big number of the world population, therefore, photographs and videos of individuals are easily usable with these networks. Face spoofing can be categorized into two groups as well. 2D face spoofing and 3D face spoofing. 2D face spoofing could be printed photo attack and replay video attack. On the other hand, 3D face spoofing could be 3D mask attack and plastic surgery attack. 2D face spoofing is performed with the help of photographs (2D images) and videos. This case of spoofing is simple to use and less costly. 3D face spoofing is performed with the help of the 3D mask, plastic surgery, and make-up. This case of spoofing needs more cost if plastic surgery is employed as a way of spoofing.

2.3 Face Anti Spoofing

Face anti spoofing is the countermeasure of face spoofing. In face recognition system, a face is accepted at sensor level and then it is passed to the face spoof detection module. If the input face is found to be a real face, then it goes to the next phase of the system or it exits if it is found to be a fake face. The objective of face anti-spoofing system is to protect the face biometric system from illicit access. In the last decade, researchers of this field introduced several face anti-spoofing systems, but it is hard to pick out one technique over the others because their operation is extremely dependent on different cases of attacks and the events they considered. A block diagram of face anti-spoofing mechanism is shown in Figure 2.

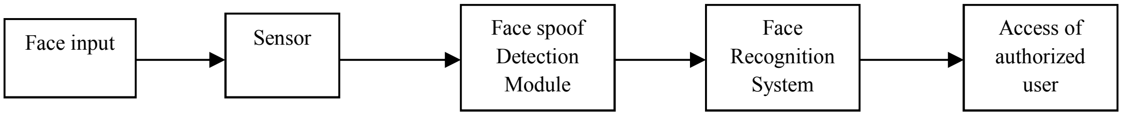

Figure 2. Anti-spoofing mechanism enabled face biometric system.

2.4 Face Spoof Attacks

The spoof attack is the action of presenting a fraud biometric testimony to a biometric system in order to attain authentication. The spoof attack uses by intruders in a face biometric system is termed as a face spoof attack. This character of attack is very simple for a face biometric system since photographs and videos of authorized users can be well captured from social network sites. Attackers try to enter into a secured system, which uses face modality for access control, by displaying printed/digital photographs or replaying recorded videos on mobile/tablet or presenting themselves after wearing a 3D mask on the face or going through a plastic surgery in front of the sensors. So, face spoof attacks can be grouped into four categories, such as (a) printed photo attack or digital photo attack, (b) replay video attack, (c) 3D mask attack and (d) plastic surgery attack.

2.4.1 Printed Photo Attack/Digital Photo Attack

When an attacker attempt to spoof a face biometric system with the help of a printed photograph of an authorized person (presenting the photograph in front of the sensor) then it is called printed photo attack. Attackers may display the photo of an authorized person on the screen of a digital device such as a tablet or mobile phone in front of the camera face, it is then called digital photo attack. Printed photo or digital photo is the most usual source of spoofing attack because anybody can download and capture the facial image of a person from social networks very well. It represents a sleazy and efficient means to do an attack.

2.4.2 Replay Video Attack

The replay video attack is performed by presenting the video of an authorized user in a digital device (i.e., laptop, mobile phone or tablet) in front of the sensor/camera of face biometric system. The video replay attack provides dynamic biometric (i.e., the motion of a user) traits. This type of onslaught is also possible because of easy accessibility of videos of a person in social sites.

2.4.3 3D Mask Attack

A mask is an object that can be made of different materials normally worn on the faces for different purposes. In 3D mask attack, attackers make a mask or clay face to spoof the system. The mask has very similar 3D face shape characteristics of the target face. The self - manufacturing of this type of mask is not an easy task. Photographic masks can be utilized for this type of approach. High resolution printed photographs, in which the eyes and the mouth portion have been cut out are used as a photographic mask. 3D face spoofing attack cost is high.

2.4.4 Plastic Surgery Attack

In our modern society due to the progress of the medical science and affordability, popularity of plastic surgery is increasing day by day. Facial plastic surgery is a component of Otolaryngology that can be split into two categories – reconstructive and cosmetic. Reconstructive facial plastic surgery corrects facial feature anomalies that may be present from birth, such as birthmarks on the face, cleft lip and palate, protruding ears, and a crooked grin. The cosmetic facial plastic surgical procedure improves the visual aspect of the facial structures and characteristics. Thus, the original facial information is altered to a great extent by this type of surgery and it is a threat to a face biometric system. After going through this type of surgery, a person can try to spoof a face biometric system and this is called plastic surgery attack.

3 Research Challenges

Identity verification using facial information is a challenging research area in computer vision. Researchers of this field face several challenges. This section enlists some of them as follows.

- Aging of subjects and complex outdoor lighting are common challenges in this area.

- It is very much difficult to detect face spoof in liveness detection based technique and motion analysis based technique by prompting the user to do some random action or challenge (such as a smiling and moving the head in a particular direction). This user's response (smiling and head movement) will provide evidence for liveness detection. The drawback of such approach is that it requires user cooperation, which makes the authentication process a time consuming. Thus making this type of approach faster is one another challenge in this field.

- The problem of recognizing a person after undergoing plastic surgery operation is still an open challenge for the automatic face authentication system.

- The face-based biometric systems may be obstructed by wearing regular makeup. Makeup may be successfully used to perform direct attacks.

- A very specific weakness is the potential for identical twins to be authorized interchangeably by a face biometric system.

- In blinking and movement of eyes based liveness detection technique, eyes glasses which causes reflection is a challenge for a face spoof detection technique.

4 Taxonomy of Face Spoof Detection Technique

In other way, the existing face spoof detection techniques [4-8[, [11], [56], [57], [60] can be divided into four groups: (a) Texture analysis based technique [5], [11], (b) Motion analysis based technique [60], (c) Liveness detection (active and passive) based technique [57] and (d) Image quality analysis based technique [60]. Figure 3 shows the categorization of face spoof detection techniques. Further, their advantages and disadvantages are given in Table 1.

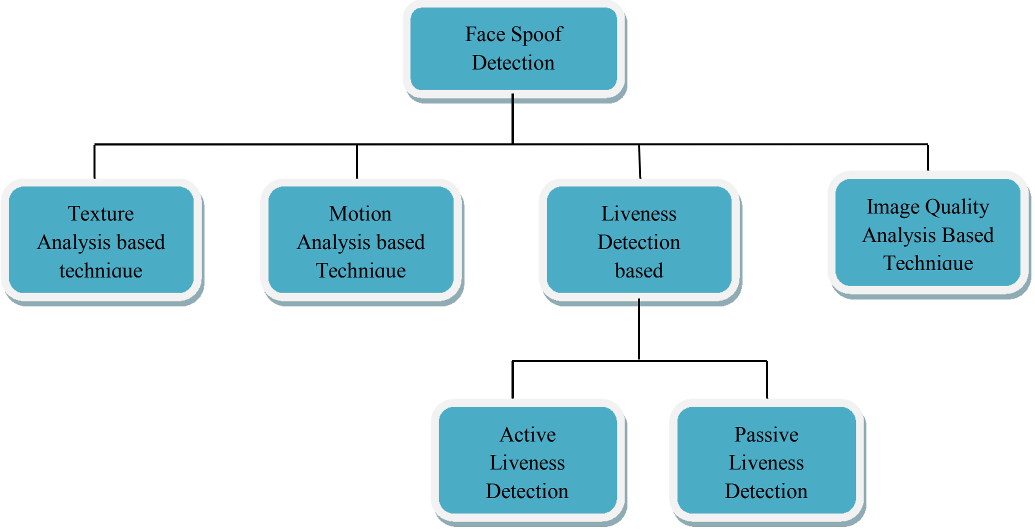
Face Spoof Detection
Texture Analysis based technique
Motion Analysis based Technique
Liveness Detection based
Image Quality Analysis Based Technique
Active Liveness Detection
Passive Liveness Detection

Figure 3. Taxonomy of Face Spoof Detection Technique

4.1 Texture analysis based Technique

Texture analysis techniques [5], [11] mainly compare the texture pattern of the face which is captured by the sensor in the system with the texture pattern of the real face which is present in the database. These techniques take the advantage of detectable texture patterns such as print failures, and overall image blur to detect attacks. This approach fights against the printed photo attack. In printed photo attack fake faces are printed on paper and presented in front of the camera for verification or identification, the printed process and the paper structure that produces texture features can differentiate those printed images from real face images. In this situation, texture analysis is useful to identify real faces because printing procedure and paper usually contain high texture characteristics. Texture analysis based approach is easy to implement and it does not need user collaboration. Textures generated by printed photographs are very diverse. Therefore, texture analysis based systems must be robust to differentiate texture patterns which require the existence of a very diverse feature set. If the presented printed photo generates low texture information, texture analysis based system does not give a good result.

4.2 Motion analysis based Technique

Motion analysis based approaches [60] mainly differentiate the motion pattern between 3D and 2D faces. Planar objects (2D faces) move significantly different from real human faces (3D objects). Motion analysis mainly depends on optical flow calculated from video sequences. Motion analysis is very hard to spoof by 2D face images. It is independent of texture and user collaboration is not needed. Motion analysis needs video, but video having low motion activity does not give good result in motion analysis. This approach can be spoofed by 3D sculptures and it needs high quality images. But Real faces display a different motion behavior compared to a spoof attempt.

4.3 Liveness detection based technique

Detecting physiological signs of life in faces those are captured by the sensor of the face biometric system is the goal of Liveness detection approach [57]. Spoofing artifacts that are used for attacks at the sensor level of a face biometric system are photographs, videos and 3D masks. Video attacks can be prevented by liveness detection. Liveness information can be gathered by user interaction with the biometric system that is captured and assessed in real time. Further, liveness detection based techniques can be categorized into two sub-groups – (a) active liveness detection and (b) passive liveness detection techniques.

Active participation of the user is needed in active liveness detection technique. User participation or even user awareness does not require in passive approach, but spontaneous eye blinks, facial expressions are used. Therefore passive anti-spoofing techniques are usually based on the detection of signs of life.

4.4 Image quality analysis based technique

The aim of this type of technique [56] is to check the different qualities of the fake faces (presented by a printed photo or digital photo) like specular reflection features, blurriness features, color diversity features, chromatic moment features, edge information and compare these qualities with the real faces. Image quality properties of real accesses and fraudulent attacks would be different. Blurriness is measured based on the difference between the original input image and its blurred version. Spoof faces are often defocused in mobile phone cameras because spoofing medium usually have limited size, and attackers have to place them close to the camera in order to conceal the boundaries of the attack medium. Due to this reason, spoof faces tend to be defocused and the image blur due to defocus can be used as a feature for anti-spoofing. Recaptured face images tend to show a different color distribution compared to colors in the genuine face images. Abnormal chromaticity in spoof faces is used as a feature to differentiate the real faces with it. The specular reflection component of a face image can be used as a feature to differentiate the fake face and real face. The color diversity tends to fade out in the spoof face due to the color reproduction loss during image recapture. Therefore color diversity is another parameter for image quality analysis based technique. Different features of the image qualities are used in this technique.

Table 1. A brief summary of different face spoof detection technique.

Name of the technique	Advantages	Disadvantages	Cost
Texture analysis based Technique	• Very simple to implement • No user collaboration needed	• Good feature vectors are needed to identify the spoof • Low quality image or video does not give good result	Low
Motion analysis based Technique	• Difficult to spoof • Low user collaboration needed	• Several video sequences are needed • Low motion activity in a video does not give good result • Low response • High computational complexity • Low robustness(can be circumvented by fake motion)	Medium
Active Liveness detection technique	• Very difficult to spoof • Cover maximum type of attacks	• High user collaboration is needed • More time consuming • Extra devices are needed	High
Passive Liveness detection technique	• Difficult to spoof • Independent from user collaboration	• Video sequences are needed	High
Image quality analysis based technique	• Very difficult to spoof • No user collaboration needed • Good generalization ability • Low computational complexity • Fast response	• High quality images are needed • Different classifiers needed for different spoof attacks.	Medium

5 Existing Face Spoof Detection Techniques

This section summarizes various approaches that are successfully used in face spoof detection. This section mainly deals with various analysis based techniques as face spoof detections and discuss their merits and demerits in addition to the basic framework of each approach.

5.1 Face Spoofing Detection using Micro-Texture Analysis

The work exploited in [11] is able to detect the face print attack with the help of a powerful feature extraction technique Multi-Scale LBP [12]. Multi-Scale LBP analyzes the texture of the facial images and encodes the micro texture patterns into an enhanced feature histogram [13]. After computing the enhanced histograms, they use a nonlinear SVM classifier with Radial Basis Function (RBF) kernel [15] to discriminate between the fake faces and real faces to prevent unauthorized access in a face biometric system. Authors use publicly available NUAA Photograph Imposter Database [14] which contains images of both real client and photo attack for their experiments. This approach uses a combination of Multi-Scale LBP operators ($LBP_{8,1}^{u2}$, $LBP_{8,2}^{u2}$, $LBP_{16,2}^{u2}$) which gives the fine details of the facial surface those are important for discriminating real human face images from fake ones. The following algorithm describes the whole work presented in [11] where LBP is used as pattern representation technique and SVM with RBF kernel is used as pattern classifier. Block diagram of the Micro-Texture analysis approach is shown in Figure 4 and various steps are described in Algorithm 1.

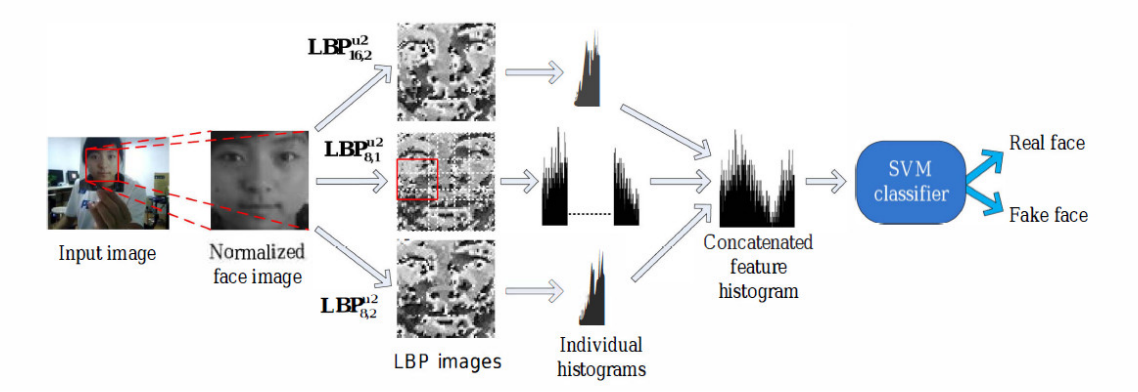

Figure 4. Block Diagram of the proposed approach [11]

Algorithm 1

Step 1: The face is first detected, then cropped and normalized into a 64×64 pixel image.

Step 2: Then $LBP_{8,1}^{u2}$ operator is applied on the normalized face image and divide the resulting LBP face image into 3×3 overlapping regions

Step 3: The local 59-bin histograms are computed and collected into a single 531-bin histogram from each region.

Step 4: Two other histograms from the whole face image using $LBP_{8,2}^{u2}$ and $LBP_{16,2}^{u2}$operators are computed, yielding 59-bin and 243-bin histograms.

Step 5: 59-bin and 243-bin histograms computed in step 4 are added to the 531-bin histogram computed in step 3. Hence, the length of the final enhanced feature histogram is 833 (i.e. 531+59+243).

Step 6: A nonlinear SVM classifier with RBF kernel is used to classify between live face and fake face using the derived feature vector.

In this experiment, the authors apply three texture operators, namely, local binary pattern (LBP), local phase quantization (LPQ) [18] and Gabor wavelets [16] to the whole facial area (without block division) separately and computed features from these operators are fed to SVM classifiers for getting the result. Equal error rate (EER) of LBP, LPQ and Gabor wavelet texture operators in discriminating live face images from fake ones are 2.9%, 4.6%, and 9.5% respectively. LibSVM Library [17] is used for SVM implementation in all experiments. Authors compare their approach with the approach proposed by Tan et. al. [14] on the same database and using the same protocol. The comparison using area under curve (AUC) method shows that this proposed approach [11] is found superior (AUC=0.99) than the approach proposed by Tan et al. (AUC=0.94). This approach is also able to achieve 98% classification accuracy, 0.6% false accept rate (FAR) and 4.4% false reject rate (FRR).

The simulation of spoofing detection takes about 16.5 ms time (in average) to process an image on a 2.4 GHz Intel Core 2 Duo CPU with 3 GB of RAM using non-optimized C++ code. This approach is robust, computationally fast and does not require user-cooperation.

5.2 Video-Based Face Spoofing Detection

This approach discussed in [20] is able to detect the video-based face spoofing attack with the help of noise signatures generated by the recaptured videos. Fourier spectrum followed by the computation of the visual rhythm [21] is used to obtain a compact representation of the captured noise and gray level co-occurrence matrices (GLCM) [24] are used as feature descriptor, since this descriptor provides spatial distribution and brightness variations of the image regions [25]. After computing the noise signature, the visual rhythm (vertical and horizontal) of each video is extracted. Extracted horizontal and vertical visual rhythms generate different texture maps. This visual rhythm is then used to classify spoof attempts in videos. They are used Support Vector Machine (SVM) classifier with two different kernels: linear and radial basis function using the LibSVM [17] and Partial Least Squares regression [23] (PLS) to discriminate the fake faces to prevent unauthorized access in a face biometric system. Authors use publicly available NUAA Photograph Imposter database [14] and Print Attack database [22] which consists of 200 videos of valid accesses of 50 different users and 200 videos of spoof attacks using printed photographs in 320×240 pixels resolutions. This approach is described in various steps in Algorithm 2. The block diagram of video-based spoof detection is shown in Figure 5.

Algorithm 2

Step 1: Extract a noise signature of every input video.

Step 2: Calculate the Fourier Spectrum on logarithmic Scale for each video.

Step 3: Extract vertical and horizontal visual rhythms for each video.

Step 4: Gray Level Co-occurrence matrices is used to extract textural information from visual rhythms.

Step 5: Extract 12 measures which summarize the textural information from each matrix: angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy and directionality.

Step 6: SVM or PLS regression method to classify the patterns that are extracted from the visual rhythms and GLCM.

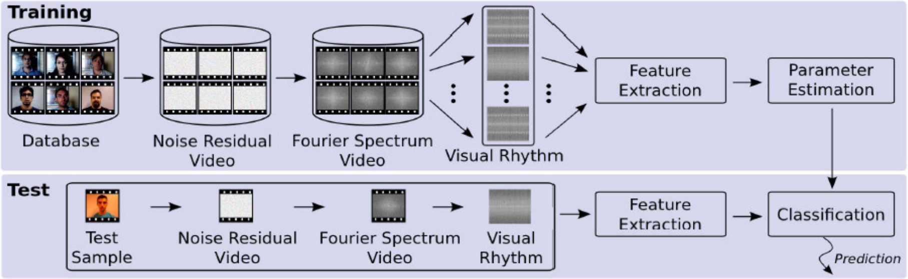

Figure 5. Block Diagram of the proposed approach [20].

In this experiment, authors use four sub-sets of their dataset: (1) Valid #1, comprising 50 valid access videos; (2) Valid #2, comprising 50 valid access videos; (3) Attack #1, with 300 attack videos created by using the monitors 1, 2 and 3; (4) Attack #2, comprising 300 attack videos created by using the monitors 4, 5 and 6. In the protocol they devised, a classifier is trained with images from a set of monitors and tested with images of monitors it never accessed to. They design two configurations for the experiments. The first configuration uses Valid #1 and Attack #1 groups to train the classifiers and Valid #2 and Attack #2 groups to evaluate the model found by the classifiers. The second configuration uses Valid #2 and Attack #2 groups to train the classifiers and Valid #1 and Attack #1 to test. The monitor characteristics are shown in Table 2.

The results show that visual rhythms calculated on a logarithmic scale Fourier spectrum represents an effective alternative to summarize videos and an important forensic signature for detecting video-based spoofs. Combining the horizontal and vertical visual rhythms produce 100% AUC in this approach. All experiments were performed on an Intel Xeon 5160, 3 GHz dual-core processor with 8GB of RAM running on Windows 7 operating system.

Table 2. Characteristics of the monitors considered in the experiments for creating the video spoofing samples.

ID	Manufacturer	Technology used in an image information	Screen type
Monitor1	Itautec	LCD	Glossy
Monitor2	LG	LCD	Matte
Monitor3	Samsung	LCD	Matte
Monitor4	LG	LCD	Matte
Monitor5	AOC	LED	Matte
Monitor6	LG	LCD	Matte

5.3 Face Spoofing Detection with Motion Magnification

This algorithm [26] is used to detect the face print attack and replay attack with the help of the powerful feature extraction techniques – Multi-scale LBP [12] and HOOF [27], and they are applied after Eulerian motion magnification approach is used to enhance the facial expression in captured video. Experiments are performed on two publicly available databases namely Print Attack Database [22] and Replay Attack Database [28]. For classification, SVM with Radial Basis Function (RBF) kernel [15] is used. The proposed approach outperforms existing approaches in terms of accuracy as well as in terms of computational time.

In this approach, at first, motion magnification is performed using Eulerian motion magnification approach [29]. Temporal intensity changes at a given position are directly amplified without the need for explicit estimation using Eulerian motion magnification approach. After performing the motion magnification on input video, face spoof detection is performed in two ways, such as (a) texture based spoof detection approach with motion magnification and (b) motion based spoof detection approach with motion magnification.

Micro and macro facial motion demonstrated by a subject are enhanced by motion magnification. This magnified motion improves the performance of spoofing detection techniques, especially texture based approaches. The motion magnification approach with LBP is presented in Figure 6 and various steps of the approach are given in Algorithm 3. Further, motion magnification with HOOF is shown in Figure 7 and various steps of the approach are given in Algorithm 4.

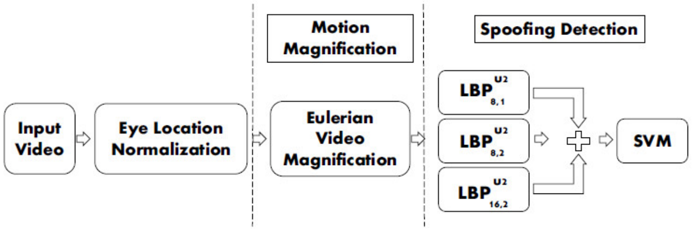

Figure 6. Block Diagram of the proposed approach [26] with motion magnification.

Algorithm 3

Step1: Eye location of face in input video is normalized.

Step 2: Eulerian Motion Magnification approach is used to enhance facial movements in the video.

Step 3: Three LBP operators ($LBP_{8,1}^{u2}$, $LBP_{8,2}^{u2}$, $LBP_{16,2}^{u2}$) with different scales are applied and three global histograms are computed.

Step 4: Concatenate three global histogram those are computed in step 3 and generate a feature descriptor of size 361(i.e. 59+59+243)

Step 5: SVM with Radial Basis Function (RBF) kernel is used to classify between real face and fake face using the feature vector which is computed in step 4.

Texture based features are widely used in different research works than motion based features. Global LBP features are efficient for spoofing detection. It is more explained in Algorithm 4.

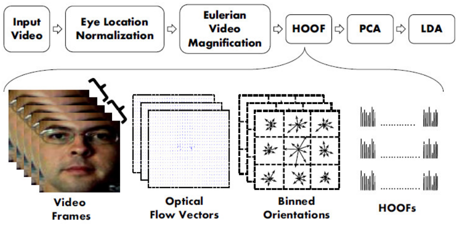

Figure 7. Block Diagram of the proposed approach [26] with motion magnification and HOOF descriptor

Algorithm 4

Step 1: Eye location of face in input video is normalized.

Step 2: Eulerian Motion Magnification approach is used to enhance facial movements in the video.

Step 3: Optical flow from the face region is computed between frames at a fixed interval.

Step 4: Histogram of the optical flow orientation angle, weighted by the magnitude is computed over local blocks and concatenated to form a single vector. This vector is called Histogram of Oriented Optical Flows (HOOF) [27].

Step 5: PCA is used to reduce the dimensionality of the feature vector.

Step 6: LDA is applied to obtain a uni-dimensional projection of the reduced feature vector and Classification is performed using thresholding and nearest neighbor approach.

The anti-spoofing feature is boosted by motion estimation using optical flow. Identification of facial micro-expressions in videos [30] is used optical flow. In this approach motion of each pixel is estimated using optical flow by solving the optimization problem [26]. Due to low computational complexity Conjugate gradient approach [31] is used to solve the optimization problem in this work.

For performance evaluation both texture and motion-based feature extraction approaches are computed with and without motion magnified videos. Experiments show that motion magnification improves the performance of LBP texture features. The HOOF descriptors obtained from motion magnified videos provide better accuracy and computational efficiency on the Print Attack and Replay Attack datasets.

The results show that only using the texture-based multiscale LBP with SVM classification yields 5.6% HTER on the Print Attack database and 11.75% on the Replay Attack database. The proposed motion magnification technique using LBP operator and SVM classification, the HTER improves to 1.25% and 6.62% on the Print Attack database and the Replay Attack database respectively. These results are better than the result produced by LBP-TOP [32] which shows 7.6% HTER on the Replay Attack database. Motion magnification approach enhances the changes in intensity values in the video, which enrich the texture of the magnified video.

The MATLAB implementations of the proposed techniques are found to be computationally efficient and time taken in execution of various stages for one video (375 frames) are – registration takes 293.8 seconds, motion magnification takes 28.4 seconds, HOOF feature extraction takes 15.2 seconds time and$LBP_{8,1}^{u2}$ +$LBP_{8,2}^{u2}$ + $LBP_{16,2}^{u2}$ feature extraction takes 14.3 seconds. The execution time of the proposed approach using HOOF feature extraction takes less time than LBP-TOP features extraction alone. It shows that the proposed approach is better than existing approaches in terms of accuracy as well as computational time.

5.4 Face Spoof Detection with Image Distortion Analysis (IDA)

A work is reported in [33] is able to detect the printed photo attack and replay video attack with the help of an Image Distortion Analysis based algorithm. This algorithm extracts four different features (specular reflection feature, blurriness feature, chromatic moment feature, and color diversity feature) from input printed photo or replay video to form the IDA feature vector. This work uses IDIAP Replay Attack Database [28], CASIA-FASD [9], and MSU MFSD for their experiment and present the result for both intra-database and cross-database scenarios. The MSU MFSD database is collected by authors using two mobile devices Google Nexus 5 and MacBook Air which contain printed photo and replayed video. For classification between a real face and a fake face, this approach uses an ensemble classifier, consisting of multiple SVM classifiers trained for printed photo and replayed video attack. The block diagram of IDA based system is shown in Figure 8 and various steps are shown in Algorithm 5.

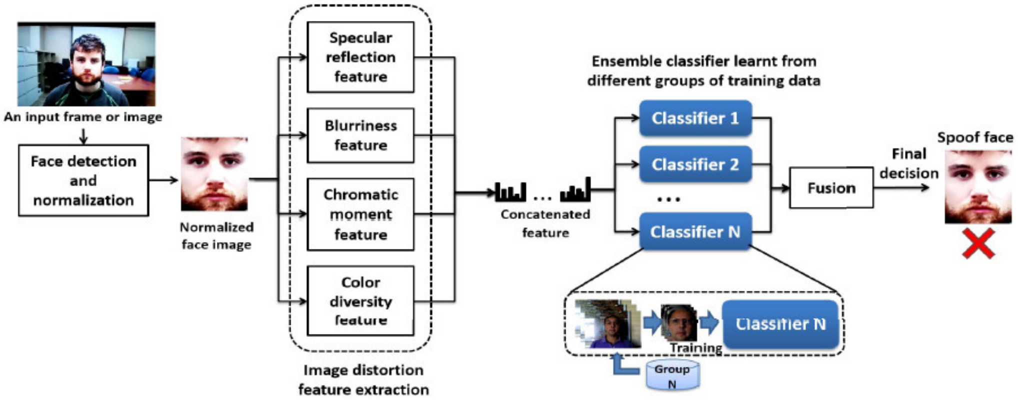

Figure 8. Block Diagram of the proposed approach [33].

Algorithm 5

Step 1: Face image is aligned based on two eye locations, and then detected face is normalized to 144×120 pixels with inter pupillary distance (IPD) of 60 pixels.

Step 2: Four different features – Specular reflection feature, Blurriness feature, Chromatic moment feature and color diversity features are extracted from normalized face forms 121-dimansional feature vector.

Step 3: Four features, extracted in step 3 are concatenated and form 121-dimensional feature vector.

Step 4: Ensemble classifier (multiple SVM) is used to classify between real face and fake face using the feature vector.

This algorithm uses the intrinsic distortions of spoof face images with respect to the genuine face images to detect the attack. For eye localization and face detection, PittPatt 5.2.2 SDK [34] is used. PittPatt 5.2.2 SDK works successfully for about 99% of the faces in the IDIAP, CASIA, and MSU face spoof databases. Face alignment and cropped face size reduces the influences of facial and background variations that are not needed to spoof detection. Finally, multiple SVM classifiers are fed by the feature vector and each classifier trained on a different group of spoof training samples, this classifier produces the final result – face input is real or fake.

5.4.1 Feature Extraction

a) Specular reflection feature: In this work, an iterative method proposed in [35] is used to separate the specular reflection component from an input face image or video frame. After calculating the specular reflection component image, specularity intensity distribution is represented by three-dimensional features: (i) specular pixel percentage *r*, (ii) mean intensity of specular pixel μ and iii) variance of specular pixel intensities *б*. The method in [35] often incorrectly classifies the mono-chromatic regions as specular components, therefore, then high-intensity mono-chromatic pixels are excluded from specular reflection component image. Pixels belonging to the intensity range (1.5μ, 4μ) are considered as specular pixels.

b) Blurriness Features: Spoof faces are often defocused due to the limited size of the spoofing medium and attackers try to cover up the boundaries of the attack medium. This cause the image blurs. Two types of blurriness features are extracted using approaches proposed in [36] and [37]. Blurriness is measured based on the difference between the original input image and its blurred version in [36]. The larger difference indicates the lower blurriness in the original image. In [37], blurriness is measured based on the average edge width in the input image. Non-reference blurriness score of these two methods is between 0~1.

c) Chromatic Moment Features: For getting Chromatic moment features, first the normalized facial image from the RGB space is converted into HSV (Hue, Saturation, and Value) space. Then mean, deviation and skewness of each channel is computed and these are considered as a chromatic feature. These three features are equivalent to three statistical moments in each channel, therefore also referred to as chromatic moment features. The percentages of pixels in the minimal and maximal histogram bins of each channel are used as two additional features.

d) Color Diversity Features: To measure the image color diversity this paper uses the method used in [38]. First, color quantization (with 32 steps in the red, green and blue channels, respectively) is performed on the normalized face image. The histogram bin counts of the top 100 most frequently appearing colors and the number of distinct colors appearing in the normalized face image are used as color diversity feature.

This work uses two testing protocols – (i) the intra-database testing protocol and (ii) cross-database testing protocol. Intra-database testing protocol uses same training set and testing set. Cross database testing protocol uses one database for training and another database for testing. This protocol provides the better generalization of the proposed approach. In CASIA spoof database, the low-quality subset (L) and normal quality subset (N) contain only long distance spoof attack samples; only the high-quality subset (H) contains short distance spoof attack samples.

a) Intra-database Spoof Detection: In this experiment, the approach is compared with LBP-TOP method proposed in [32], the DoG baseline method proposed in [40] and two baseline methods LBP+SVM [11] and DoG-LBP+SVM [42] using IDIAP, CASIA (H protocol), and the MSU database. The experiment shows that proposed IDA+SVM method outperforms the other methods in most intra-database scenarios. This approach shows better (per frame) HTER and ROC than LBP-TOP on IDIAP. The performance of the proposed approach is close to the two best methods LBP-TOP and LBP+SVM on CASIA (H protocol). The proposed approach achieves 94.7% TPR @ 0.1 FAR and 82.9% TPR @0.01 FAR which is much better than those baseline methods when 35 subjects from MSU database are used for training. This experiment shows that proposed IDA feature have the similar discriminative ability for intra-database testing like LBP and LBP-TOP which have an excellent discriminative ability to distinguish a fake face from a real face. One another baseline method image quality analysis (IQA) [43] shows 15.2% HTER on IDIAP replay attack database, but the proposed method achieve 7.4% HTER on the same database.

b) Cross-database Spoof Detection: In this experiment, two groups of cross-database performance were evaluated: (i) IDIAP versus MSU and vice versa, and (ii) CASIA (H) versus MSU and vice versa. IDA features outperform LBP and DoG-LBP features in cross-database testing. IDA features are more robust, for both replay attack samples and printed attack samples, in almost every cross database testing scenario. For the replay attack samples, the IDA features achieve average 90.5% TPR @ 0.01 FAR and for the printed attack samples, the IDA features achieve an average 31.2% TPR @ 0.01 FAR when trained on IDIAP and tested on MSU, which is still much better than the LBP and DoG+LBP features. IDA features show better cross-database performance in the replay attack samples compared to the printed attack samples. IDA features show better cross-database performance between similar cameras.

The total time consumption for face normalization, feature extraction, and classification (including the two constituent classifiers) is about 0.26s per frame on a testing platform with Intel(R) Core(TM) i7-4700MQ CPU @ 2.40 GHz and 8GB RAM, the time needed for face detection is not considered here.

5.5 Live Face Video vs. Spoof Face Video

Another novel approach presented in [44] is able to detect the spoof face video with the help of moiré pattern which is appear when a video is recaptured or photo replays on a screen. The characteristics of moiré pattern are represented with the help of Multi-scale LBP and DSIFT features. IDIAP replay-attack database, CASIA database [9] and a database collected in their own laboratory (RAFS), which is based on the MSU-MFSD database are used for their experiments. The proposed approach is very effective in face spoof detection for both cross-database and intra-database testing scenarios. The domain of interest of this research lies in mobile phone unlock. Figure 9 shows the block diagram of the approach exploited in [44] and various steps are shown in Algorithm 6.

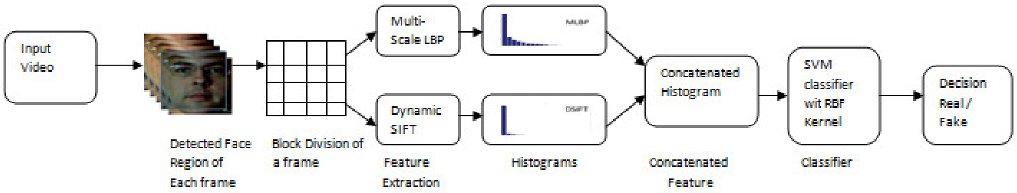

Figure 9. Block Diagram of the proposed approach.

Undesired aliasing of images are known as moiré pattern, those are produced during various image display and image acquisition processes [45]. When reconstructed signal does not well represent the original signal, this effect called aliasing. When two or more patterns are overlaid on top of each other and generate a third new pattern this is called moiré pattern. Moiré patterns are inexorable in color printing with CMYK half-toning model and also observed in the screen shooting photography. Real world scenes are not accurately represented by images with moiré patterns. Spatial frequency differences between the display and acquisition devices are the main reason to produce the moiré pattern in screen shooting photography.

In experiment after analyzing the 310 replay attack videos from three different databases, different moiré patterns are generated across the replay attack video frames. MLBP [12] and SIFT [41] texture descriptors are used to capture moiré pattern in this proposed approach.

Algorithm 6

Step 1: Each video is decoded into individual frames. (Using the FFmpeg library)

Step 2: Face or Face region is detected from input frame. (This step is optional)

Step 3: Input frame is first divided into 32 X 32 patches with an overlap of 16 pixels between every two successive patches.

Step 4: For each image patch Multi-scale LBP is applied to capture moiré pattern.

Step 5: Densely sampled SIFT (DSIFT) features from each image patch are calculated using 8 Orientation bins and 16 segments

Step 6: SVM classifier with a RBF kernel is used for classification between real face and fake face using histogram bins of Multi-scale LBP, DSIFT, and concatenation of Multi-scale LBP and DSIFT.

Authors use Multi-scale LBP to capture the characteristics of moiré patterns for each image patch. This proposed approach also uses densely sampled SIFT (DSIFT) for feature extraction. Without performing the face detection operation, spoof face detection is possible in this approach because moiré patterns exist in the whole video frame, not only in faces. The robustness of this approach is evaluated by (i) whole video frame, (ii) detected face image, and (iii) bottom part of the face image. In classification stage one voting scheme is used which based on multiple frames. In a video if more than 50% frames are fake, then it indicates that the video is fake otherwise the video is live.

Two testing protocols – (i) cross-database testing and (ii) intra-database testing are applied on IDIAP, CASIA and RAFS databases to evaluate the performance of the proposed approach. In cross-database testing, classification accuracy on IDIAP database is 88.0%, on CASIA database is 67.0% and on RAFS database is 85.5%. In cross-database testing HTER of the proposed method is 18.0% on the IDIAP, 49.0% on CASIA, and 11.4% on RAFS database which better than [3], which show 47.1% HTER and 48.9% HTER on the IDIAP and CASIA database respectively under cross-database testing. In this approach, MLBP feature extraction takes 0.09 seconds per frame, and classification takes 0.02 seconds per frame with an MATLAB implementation on a Windows 7 platform with Intel Core 2 quad3.0 GHz CPU and 8 GB RAM.

In intra-database testing, the proposed approach achieves 3.3%, 0.0% and 11.3% HTER on the IDIAP, CASIA, and RAFS database respectively. HTER achieved on CASIA database in this approach, is smaller than the HTER (11.8%) achieved by [19].

5.6 Face spoofing detection from single images using texture and local shape analysis

Face images captured from printed photos may visually look very akin to the images captured from live faces. The proposed approach [5] design a powerful feature space which has the eptitude to differentiate between images of real face and face prints. This approach explores two texture based (LBP [12] and Gabor wavelet [16]) and one gradient based (HOG [47]) feature extraction technique to detect whether there is a live person or fake face in front of the camera. LBP operator and Gabor filters are used to extract textural features (micro-texture pattern and macroscopic information) from face images. Histogram of oriented gradients (HOG) is a gradient based local shape descriptor which provides additional information to the face description. Using three different face descriptors (LBP, Gabor Wavelet and HOG), three face representations are generated and to transform the data into compact linear representation, a homogeneous kernel map [48] is applied on each resulting feature vector. The fast linear Support Vector Machine (SVM) [15] is then applied on each transformed feature space and the score level fusion of the individual SVM outputs are used for final decision, whether the input face is real or fake. NUAA Photograph Imposter Database [14], Yale Recaptured Database [50] and Print-Attack Database [22] are used to evaluate the performance of the proposed technique. When experiments are conducted on NUAA Imposter face database, block size of 21×21 pixels was used for extracting HOG features. When experiments are conducted on Yale Recaptured face database, block size of 8×8 pixels was used for extracting HOG features.

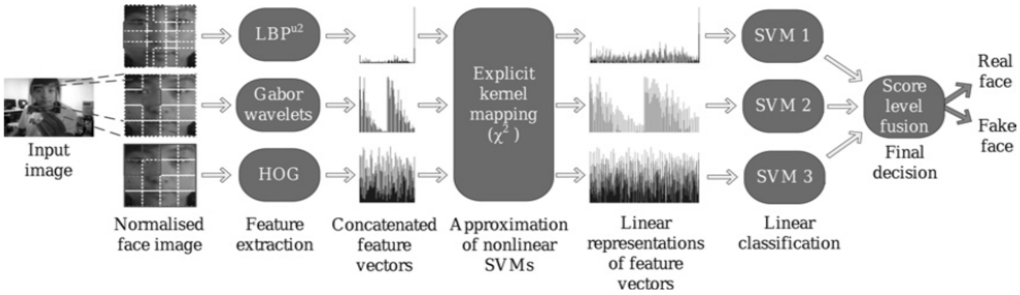

Figure 10. Block Diagram of the proposed approach [5].

Algorithm 7

Step 1: The face is detected, cropped and normalized into an M×M pixel image.

Step 2: Facial images are spatially partitioned into several regions.

Step 3: LBP, Gabor wavelet and HOG face representations are computed separately using different block divisions.

3a: $LBP_{8,1}^{u2}$ is applied on 3×3 overlapping regions (with an overlapping size of 14 pixels) of face images to extract the spatial feature and generate the concatenated histogram. $LBP_{8,2}^{u2}$ and $LBP_{16,2}^{u2}$ are also applied on whole face image and compute two global LBP histograms.

3b: Gabor filter is applied on 4×4 equally spaced non-overlapping regions and extract 40 Gabor wavelets of 5 different scales and 8 orientations.

3c: HOG feature are extracted from 8 orientations.

Step 4: A homogeneous kernel map is applied on each face description and generate their corresponding linear approximation of a X^2 kernel.

Step 5: Feature representation produced by step4 is fed to a fast linear Support Vector Machine (SVM) classifier to generate the score.

Step 6: Z-score normalization technique and weighted score level fusion are applied to combine the outputs of individual SVM for decision (input image is a live face or not).

The block diagram of the approach discussed in Section 5.6 is given in Figure 11. Various steps are described in Algorithm 7. The combination of all three features (LBP, Gabor wavelet and HOG) leads to excellent results. At first experiments are conducted on the NUAA Photograph Imposter Database and compare the result of proposed technique with the best results in [14, 50, 11] using the same protocol. The comparative results reveals that the proposed approach gives an improvement in Equal Error Rate (EER) from 2.8% to 1.1% and in area under curve (AUC) from 0.995 to 0.999. Authors validate the use of homogeneous kernel map by comparing it to non-linear SVM used in [11]. Then experiments are conducted on Yale Recaptured Database and Print-Attack Database and compare the performance of proposed approach to previously published results using these two databases. All experiments on Yale Recaptured Database have been performed using ten fold cross validation like in [50]. Result of the proposed approach on Yale Recaptured Database compared with the results of [50, 11] and the proposed approach gives excellent result (100% accuracy, 100% positive rate and 0.0% False Positive Rate). Finally experiments are conducted on Print Attack Database. OpenCV library implementation of the Viola-Jones algorithm [51] was used for face detection and 2D cascaded AdaBoost [52] is used to retrieved eye locations. The face images are geometrically normalized according to the detected eye coordinates and cropped into grayscale images of 80×80 pixels. The overlapping size of LBP blocks of 17 pixels and block size of 10×10 pixels was used for extracting HOG features while experiments are conducted on Print-Attack Database. The performance of proposed approach compared to the approaches presented by the teams who participated in IJCB 2011 competition on counter measures to 2D facial spoofing Attacks [53]. The proposed approach gives 1.67% FAR, 1.67% FRR on Development set and 0% FAR, 0% FRR and 0% HTER on test set. The Linear SVM implementation of LIBLINEAR [54] and a three-dimensional approximated feature map computed with VLFeat [55] are used in all experiments.

The proposed approach is robust, computationally fast and does not require user cooperation. The texture features that are used for face spoof detection can also be used for face recognition. This provides a unique feature space for coupling spoof detection and face recognition. The block diagram of this approach is shown in Figure 10.

5.7 Face Anti-Spoofing Based on General Image Quality Assessment

The image quality of human faces and face prints shows different characteristics because a human face is a complex non-rigid 3D object where as a photograph is a planar rigid object. This work introduced in [56], is an Image Quality Analysis based approach used to discriminate between legitimate and imposter access in a face biometric system. This approach is able to detect replay video attack and photo attack. 14 image quality features are banished from one image to distinguish between legitimate and imposter samples. Figure 11 shows the block diagram of the GIQA based approach and various steps are described in Algorithm 8.

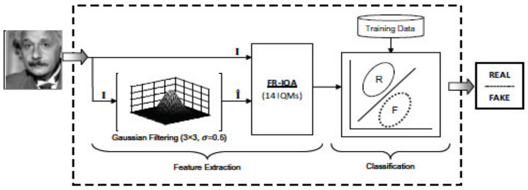

Figure 11. Block Diagram of the proposed approach [56].

Algorithm 8
Step 1: Generate a distorted version *Î* of the input gray-scale image *I* with a low-pass Gaussian filter (σ =0.5 and size 3 × 3). **Step 2:** 14 image quality measures (MSE, PSNR, SNR, SC, MD, AD, NAE, RAMD, LMSE, NXC, MAS, MAMS,TED and TCD) are computed from both images *I* and *Î* and derive Quantitative scores of each image quality measure. The feature vector is derived from Quantitative scores of each image quality measure. **Step 3:** Simple Linear Discriminant Analysis (LDA) is used to classify real/fake faces.

This algorithm employs 9 pixels wise difference measures – Mean Squared Error (MSE), Peak Signal to Noise Ratio (PSNR), Signal to Noise Ratio (SNR), Structural Content (SC), Maximum Difference (MD), Average Difference (AD), Normalized Absolute Error (NAE), R-Averaged Maximum Difference (RAMD) and Laplacian Mean Squared Error (LMSE); 3 correlation based measures - Normalized Cross-Correlation (NXC), Mean Angle Similarity(MAS) and Mean Angle Magnitude Similarity (MAMS); and 2 edge related quality measures – Total Edge Difference (TED) and Total Corner Difference (TCD). The Sobel operator is used to build the binary edge maps I_E and $\hat{I}_E$. The Harris Corner detector is used to compute the number of corners found in *I* and *Î*. Different image quality measures are enlisted in Table 3.

Table 3. Image quality measures [56] are shown.

#	Name of the IQMs and Reference	Description
1	Mean Squared Error (MSE) [74]	MSE(I, $\hat{\mathrm{I}}$) = $\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(\mathrm{I}_{i,j}-\hat{\mathrm{I}}_{i,j})^2$
2	Peak Signal to Noise Ratio (PSNR) [69]	PSNR(I, $\hat{\mathrm{I}}$) =10log($\frac{\max(\mathrm{I})^2}{MSE(\mathrm{I}-\hat{\mathrm{I}})}$)
3	Signal to Noise Ratio (SNR) [58]	SNR(I, $\hat{\mathrm{I}}$) =10log($\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(\mathrm{I}_{i,j})^2}{N\cdot M\cdot MSE(\mathrm{I},\hat{\mathrm{I}})}$)
4	Structural Content (SC) [25]	SC(I, $\hat{\mathrm{I}}$) = $\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(\mathrm{I}_{i,j})^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}(\hat{\mathrm{I}}_{i,j})^2}$
5	Maximum Difference(MD) [25]	MD(I, $\hat{\mathrm{I}}$) = $\left\vert\mathrm{I}_{i,j}-\hat{\mathrm{I}}_{i,j}\right\vert$
6	Average Difference (AD) [25]	AD(I, $\hat{\mathrm{I}}$) = $\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(\mathrm{I}_{i,j}-\hat{\mathrm{I}}_{i,j})$
7	Normalized Absolute Error (NAE) [25]	NAE(I, $\hat{\mathrm{I}}$) = $\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left\vert\mathrm{I}_{i,j}-\hat{\mathrm{I}}_{i,j}\right\vert}{\sum_{i=1}^{N}\sum_{j=1}^{M}\left\vert\mathrm{I}_{i,j}\right\vert}$

8	R-Averaged MD (RAMD) [74]	RAMD(I, $\hat{\mathrm{I}}$,R) = $\frac{1}{R}\sum_{r=1}^{R}\max_r\left\|\mathrm{I}_{\mathrm{i,j}}-\hat{\mathrm{I}}_{\mathrm{i,j}}\right\|$
9	Laplacian MSE (LMSE) [25]	LMSE(I, $\hat{\mathrm{I}}$) = $\frac{\sum_{i=1}^{N-1}\sum_{j=2}^{M-1}(h(\mathrm{I}_{\mathrm{i,j}})-h(\hat{\mathrm{I}}_{\mathrm{i,j}}))^2}{\sum_{i=1}^{N-1}\sum_{j=2}^{M-1}h(\mathrm{I}_{\mathrm{i,j}})^2}$
10	Normalized Cross-Correlation (NXC) [25]	NXC(I, $\hat{\mathrm{I}}$) = $\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(\mathrm{I}_{\mathrm{i,j}}\cdot\hat{\mathrm{I}}_{\mathrm{i,j}})}{\sum_{i=1}^{N}\sum_{j=1}^{M}(\mathrm{I}_{\mathrm{i,j}})^2}$
11	Mean Angle Similarity (MAS) [74]	MAS(I, $\hat{\mathrm{I}}$) =1- $\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(\frac{2}{\pi}\cos^{-1}\frac{\langle\mathrm{I}_{\mathrm{i,j}},\hat{\mathrm{I}}_{\mathrm{i,j}}\rangle}{\Vert\mathrm{I}_{\mathrm{i,j}}\Vert\Vert\hat{\mathrm{I}}_{\mathrm{i,j}}\Vert})$
12	Mean Angle Magnitude Similarity (MAMS) [74]	MAMS(I, $\hat{\mathrm{I}}$) = $\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(1-[1-\alpha_{i,j}]\left[1-\frac{\Vert\mathrm{I}_{\mathrm{i,j}}-\hat{\mathrm{I}}_{\mathrm{i,j}}\Vert}{255}\right]\right)$
13	Total Edge Difference (TED) [32]	TED(I, $\hat{\mathrm{I}}$) = $\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left\|\mathrm{I}_{\mathrm{E}_{\mathrm{i,j}}}-\hat{\mathrm{I}}_{\mathrm{E}_{\mathrm{i,j}}}\right\|$
14	Total Corner Difference (TCD) [32]	TCD(N_{cr}-$N\hat{}_{cr}$) = $\frac{\left\|N_{cr}-N\hat{}_{cr}\right\|}{\max(N_{cr},N\hat{}_{cr})}$

In the RAMD entry in Table 3, max_r is the r-highest pixel difference between two images. For the present implementation, R = 10. In the LMSE entry in Table $h(I_{i,j}) = I_{i+1,j} + I_{i-1,j} + I_{i,j+1} + I_{i,j-1} - 4\, I_{i,j}$. In the MAMS entry in the Table $\alpha_{i,j}$ = $\frac{2}{\pi}\cos^{-1}\frac{\langle\mathrm{I}_{\mathrm{i,j}},\hat{\mathrm{I}}_{\mathrm{i,j}}\rangle}{\Vert\mathrm{I}_{\mathrm{i,j}}\Vert\Vert\hat{\mathrm{I}}_{\mathrm{i,j}}\Vert}$. In the TCD entry in the table N_{cr} and $N^{\wedge}_{cr}$ are the number of corners found in I and $\hat{I}$.

The approach is tested on publicly available Replay Attack DB [28] and CASIA-FAS DB [9], which contain both photo and video attacks. The proposed image quality analysis based method is a single image technique (it needs one input image and not a sequence of them) and each frame of the videos in the REPLAY-ATTACK database has been considered as an independent sample. Therefore, classification (real face or fake) is done on a frame-by-frame basis and not per video. The hand-based mobile photos give 2.6% HTER, hand-based print photos give 9.3% HTER and hand-based highdef photos give 13.1% HTER on Replay-Attack database. The fixed-support mobile photos give 3.5% HTER, fixed-support print photos give 8.4% HTER and fixed-support highdef photos give 9.1% HTER on Replay-Attack database. The low-resolution warped photos give 25.0% HTER, low-resolution cut photos give 23.3% HTER and low-resolution videos give 21.7% HTER on CASIA-FAS DB. The normal-resolution warped photos give 23.3% HTER, normal-resolution cut photos give16.7% HTER and normal-resolution videos give 23.3% HTER on CASIA-FAS DB.

The high-resolution warped photos give10.0% HTER, high-resolution cut photos give 11.7% HTER and high-resolution videos give 6.7% HTER on CASIA-FAS DB.

5.8 Face Liveness Detection for Combating the Spoofing Attack in Face Recognition

Typical counter measure against spoofing is liveness detection and this aims in detecting physiological signs of life such as eye blinking, facial expression changes, mouth movements and so on. To increase the security of the face recognition system, face liveness detection is very important to distinguish the images captured from a live face from a forged face. The technique proposed in [57] is a face liveness detection method used Dynamic High Frequency Descriptor (DHFD) based on the High Frequency Descriptor (HFD) [58] to combat the spoofing attack. The proposed approach requires a flash as the additional illumination. Two consecutive images, including one with and another without flashlight, are captured from the subject. The High Frequency Descriptor (HFD) values of both images are used to detect the face liveness. The additional illumination increase the details of the skin and hair of a real person, but it causes a strong glisten and lower the High Frequency Descriptor (HFD) of a spoofing screen. Hence the High Frequency Descriptor (HFD) values of a real and fake images are different. The proposed HFD based liveness detection method confront the high resolution planar media applied in the attack. The proposed approach not only calculate the HFD value, it also calculate DHFD value which is capable of counterbalancing the effect of high resolution attack by evaluating the difference of the high frequency components between two images. Authors create their own database and used that in the experiment because the proposed technique require an additional flash as an auxiliary tool, but publicly available databases for face liveness detection, e.g. Print-Attack database [22], Replay-Attack database [28], cannot be used. Figure 12 shows the block diagram of the system and Algorithm 9 shows the various steps of the system.

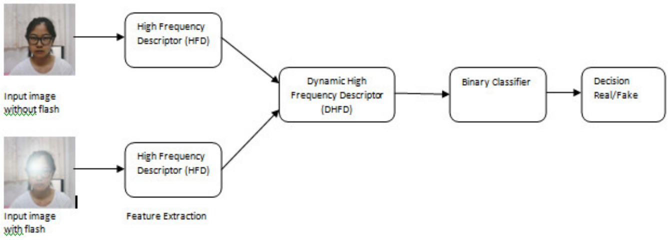
Input image without flash
High Frequency Descriptor (HFD)
Dynamic High Frequency Descriptor (DHFD)
Binary Classifier
Decision Real/Fake
Input image with flash
High Frequency Descriptor (HFD)
Feature Extraction

Figure 12. Block Diagram of the proposed approach

Algorithm 9

Step 1: Two consecutive images are taken for each user in a very short interval. The former image is taken without Flashing while the latter one with flash.

Step 2: The Dynamic High Frequency Descriptor (DHFD) of these two images is calculated as

$$DHFD = \frac{HFD_f - HFD_l}{HFD_f} \quad \text{where}$$

$$HFD = \frac{\iint_{\Omega=\{(u,v)|\sqrt{u^2+v^2}>r.f_{\max} and |F(u,v)|>T_f\}} |F(u,v)|\, dudv}{\iint |F(u,v)| dudv - F(0,0)}$$

Step 3: Take decision based on DHFD value.

If DHFD value is positive, face image is fake.

If DHFD value is negative, face image is real.

In above mentioned algorithm *F(u, v)* is the Fourier transform of the facial image. f_{max} is the highest radius frequency of *F(u, v)*. *r* defines the range of high frequency and T_f is the threshold for frequency magnitude. This method is evaluated on authors own database under spoofing attack with a high resolution screen. Authors collect 42 facial samples in their dataset, 21 samples from live faces and other 21 samples are from fake faces. Each sample consists of two consecutively taken images, one with flash and other one without flash. The fake samples are simulated by first displaying the flash-free live face images on a 24-inch full HD LCD display and then recapturing these images. All the images are taken using a Canon EOS 600D DSLR camera with the parameters fixed on 1/3 sec exposure, f/5.6, ISO 200, and 55mm focal length. The entire dataset is preprocessed before being applied Fourier transform. All the images are converted from RGB image to gray scale intensity image. Then face region is detected using Local SMQT features and split up Snow classifier [59] [11]. Finally, the entire dataset is duplicated twice and normalized to 1000×1000, 500×500, and 250×250 pixels respectively to imitate attacks under different capture resolution. *r* is set to 0.2, 0.2, 0.3, T_f is set to 5000, 10000, 30000 and Thfd is set to 110, 35, 5 respectively for experiments. The experimental results shows that the proposed approach retains high viability even when the capture resolution is high. The proposed approach gives 95% precision (is the fraction of retrieved positive samples that are relevant) when the resolution is 500×500 and 250×250 and 85.71% precision when the resolution is 1000×1000. The proposed technique performs outstanding compared to HFD.

5.9 Motion Based Counter-Measure to Photo Attacks in Face Recognition

The movement of planar objects (video display and photographs) differs significantly from real human faces which are complex 3D objects. When real faces move, the background motion is not present there, but in case of video display or photograph, both facial motion and background motion is found to be same. The technique presented in [60] is based on foreground/background motion correlation using optical flow. It tries to detect motion correlations between the head of the user trying to authenticate and the background of the scene. This technique uses a video as input and converts the input video to gray scale, and then optical flow is computed. For experiments, authors used the Liu's [61] implementation for estimating dense optical flow, as it is developed in C++, it executes in much faster pace and could be easily ported into their unified framework. The core of the algorithm in this implementation is based on [62], [63]. All experiments are conducted on Photo-Attack Database [22]. It outperforms all other algorithms achieving nearly perfect scoring with an equal error rate (EER) of 1.52% on the available test data. The block diagram of the system is shown in Figure 13 and the Algorithm 10 describes the various steps used in designing the system.

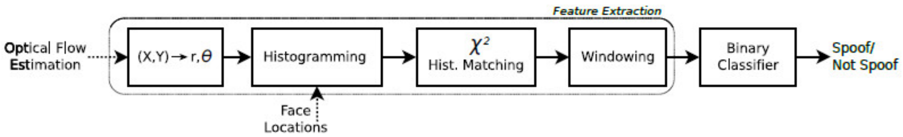

Figure 13. Block Diagram of the proposed approach [60]

```
Algorithm 10
Step 1: Take a video as input
Step 2: Convert the input video to gray scale
Step 3: Extract features
    3a: Compute the direction $\Theta$ of motion for every pixel using the horizontal and vertical orientations according
        to a simple Cartesian to polar coordinate transformation. (Discard the magnitude and preserve only the
        movement direction of every point)
            $\Theta_{i,j} := atan2(V_{ij},U_{ij})$
    3b: Calculates the normalized histograms for face and background regions based on the quantized angle for
        every flow field frame.
    3c: Calculate the $X^2$ distance between the angle histogram of face and background regions of each frame.
    3d: Average the $X^2$ scores over a window size of $N$ frames.
Step 4: Scores generated in 3d is fed to the binary classifier to take the decision.
        If scores close to 0, attacks are expected
        Else face is real
```

The input consists of the horizontal and vertical velocity estimates and it also uses the face bounding boxes available in the database to separate the features from face and background regions. First the direction (ϑ) of motion for every pixel using the horizontal and vertical orientations are computed according to a simple Cartesian to polar coordinate transformation. The movement of direction of every pixel is considered and the magnitude components are discarded. The histogram computation unit calculates the normalized histograms for face and background regions. The normalized histograms are based on the quantized angle for every flow field frame. The face locations for every frame are supplied as input to this unit. The number of bins (Q) at the angle histogram is a hyper-parameter of this algorithm and the offset used for the first bin starting angle. In the same way, this algorithm can be tuned to capture different kinds of movements.

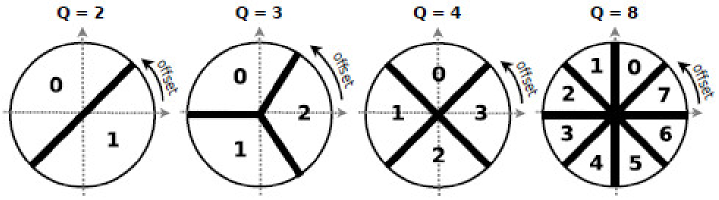

Figure 14. Offsetting for the proposed approach [60], if considering a different number of bins (Q) for the histogramming block.

After histogram normalization, X^2 distance between the angle histograms of face and background regions are calculated per frame. If F_i is a bin value at angle histogram of a face region, B_i is the corresponding bin value in the angle histogram of the background region and i is the bin number, then X2 distance between the angle histograms of face and background region is as follows:

$$X^2(F,B)=\sum_i \frac{(F_i-B_i)^2}{(F_i+B_i)}$$

The windowing unit averages the X2 scores over a window size of N frames. The scores computed from the windowing unit are fed to the binary classifier, which detects the spoof attacks based on a threshold on the equal error rate (EER) tuned at the development set. If the scores are close to 0 then attacks are expected because the motions of the face and the background area are same in the case of attack. Scores of real accesses should be greater than 0 due to the fact that the face region moves independently from the background region. To deploy the proposed algorithm, tuning of three hyper-parameters are needed. *Q* = 2 up to *Q* = 8 are used at the histogramming quantization unit (Figure 14). Authors fixed the offset used so bin 0 starts parallel to the vector (x, y) = (1, 0). N is set to 220 frames with an overlap of 219 (N - 1) frames for windowing unit. The experimental result gives EER of 0.36% and HTER of 1.52% with an Overlap of 219 frames while full background size is considered, if background size is reduced to 50% then this motion-based approach gives EER of 0.06% and HTER of 1.76%.

6 Competitions Results on Face Spoof Detection

Now a days, the face recognition systems are widely used in such environments where security is uncompromised, nevertheless face spoofing attack is one of the serious problems which decrease the reliability of face recognition system. Thus, to protect the authentication process, face recognition systems must be able to reject the use of photograph or video face image instead of using the live face. However, the number of anti-spoofing systems is still limited in use. Different researchers of this field presented different state-of-the-art algorithms, using different databases and different evaluation protocols. Therefore, a direct comparison between these approaches is not possible. To overcome this problem, in year the 2011 IJCB [53] organized one competition with a goal to compare and evaluate anti-spoofing algorithms using a common database (Print-Attack database [22]) with a unique protocol. This competition only addresses the photo attack. Day by day, more and more advanced face spoofing attacks are introduced by the intruders. In year the 2013 Tabula Rasa organized 2nd competition [49] for the comparison of the performance of spoofing countermeasures on Replay-Attack database [28] using a unique evaluation protocol to detect printed photo attacks, replay photo attacks and replay video attacks.

6.1 Competition on Counter Measure to 2-D Facial Spoofing Attacks

Facial recognition techniques are currently deployed in a wide range of applications for access control, security, law-enforcement and database indexing, etc. Facial recognition has the advantage of non-intrusiveness over the other biometrics system such as irises and fingerprints. As in any other field of Biometrics such as speaker or fingerprint verification, 2-D facial recognition systems are subject to attacks. Spoofing identities using photographs are one of the most usual techniques to attack 2-D face recognition. Before 2011 there was no comparative study of different techniques of face anti-spoofing which use the same protocols and data. International Joint Conference on Biometrics 2011 (11th -13th October'2011, Washington DC, USA) organized a competition [53] on Counter Measure to 2-D Facial Spoofing Attacks. The motivation behind this competition was to compare the operation and performance of different face anti-spoofing algorithms on the same database using a unique evaluation method. Six different teams from universities around the world have participated in the contest. In this competition, a common trend was to use multiple anti-spoofing measures combining motion, liveness and texture and the participants were able to achieve impressive results. The results suggest the investigation of more complex attacks. Performance comparison between the teams has been summarized in Table 4.

Teams: Six teams those had participated in the competition are:

i) Ambient Intelligence Laboratory (AMILAB), Italy

ii) Center for Biometrics and Security Research, Institute of Automation, Chinese Academy of Sciences (CASIA), China

iii) Idiap Research Institute (IDIAP), Switzerland

iv) Universidad de Las Palmas de Gran Canaria, (SIANI), Spain

v) Institute of Computing (UNICAMP), Brazil

vi) Machine Vision Group (UOULU), University of Oulu, Finland.

Database and Protocol used:

Publicly available PRINT-ATTACK database [22] was used to evaluate the performance of the algorithms presented by six participating teams. In competition, initially training and development data sets were given to contestants. Training set contains 60 real accesses and development set also contains 60 attack attempts. All the teams were given a couple of months to train and develop their classification system. The test data set contains 80 real access and 80 attack videos. All videos in the test set contained 230 frames. To evaluate the performance of the algorithms properly, files of the test data set are anonymized to conceal the type of the video (real/attack). Organizers had asked all the teams to provide scores of the development set and test set. Organizers of this competition were not interested about speed, latency and complexity of the contestant's methods, they only interested about the performance of the algorithms in terms of spoof-detection. To compute the performance measure of the spoofing detection systems, at first a threshold at Equal Error Rate (EER) on the development set scores is computed. Then, on the test set scores using the same threshold, they computed Half Total Error Rate (HTER), which combines the False Rejection Ratio (FRR) and the False Acceptance Ratio (FAR) with 0.5 weight. Spoof detection accuracy is good if FAR/FRR/HTER value is close to 0%.

Methods:

1) AMILAB –

This team extracts color features, edges and texture features from printed photos and real scenes, and used a set of Support Vector Machines (SVMs) to compute a frame level confidence score of being a real session or not. To obtain a high separation between scores distributions, this team combined similarity scores by means of the Dynamic Score Combination methodology [64].

2) CASIA –

The method presented by this team is based on three observations –

i) Real access videos tend to have non-rigid motions, especially in the eye and mouth regions, while printed photos only have rigid transformations like translation, scaling and rotation.

ii) Real access videos tend to have less noise than those spoofing videos.

iii) Real access videos only have local motions in the face region while spoofing videos usually have global motions spread-out the whole support.

The first observation is tested using a batch image alignment technique proposed in [65], called "RASL". Observation 2 is tested using noise variance [66] and [67]. Observation 3 is tested using ratio of motion between face region and background.

3) UNICAMP –

This team integrates feature descriptors based on histogram of oriented gradients (HOG) [47], grey level Co-occurrence matrix (GLCM) [24], and histograms of shearlet coefficients (HSC) [68] with a weighting scheme based on partial least sequence regression [69].

4) IDIAP –

This scheme processes each video by accumulating a single Local Binary Pattern (LBP) code histogram [12] with data from every single image in the stream. The histogram is matched against a pre-calculated model, using the X^2 method as proposed on the same reference to generate a final score.

5) SIANI –

This team used Chow and Liu algorithm [70]. The classification of videos between real and attacks is done with the Bayesian Network approach [71] included in the Weka open source software [72].

6) UOULU –

This team used Sobel horizontal edge emphasizing filter to highlight the image defects and produce a gradient image from which a single Local Binary Pattern (LBP) feature histogram is computed [12]. The 59 bin histogram is fed to an SVM classifier which determines whether or not the window contains a real face.

Table 4. Techniques used by the teams and performance comparison between the teams who participated in the IJCB 2011 competition on counter measure to 2D facial spoofing attacks.

Method	Techniques used	Development set		Test set		
		FAR	FRR	FAR	FRR	HTER
AMILAB	Motion Analysis, Texture Analysis, Liveness Detection	0.00	0.00	0.00	1.25	0.63
CASIA	Motion Analysis, Texture Analysis	1.67	1.67	0.00	0.00	0.00
IDIAP	Texture Analysis	0.00	0.00	0.00	0.00	0.00
SIANI	Motion Analysis	1.67	1.67	0.00	21.25	10.63
UNICAMP	Motion Analysis, Texture Analysis, Liveness Detection	0.00	0.00	1.25	0.00	0.00
UOULU	Texture Analysis	1.67	1.67	0.00	0.00	0.00

6.2 The Second Competition on Counter Measures to 2D Face Spoofing Attacks

Spoofing attacks are a big threat to biometric systems when high security is required. There is multiple ways (photo print of valid user, video playback or 3D mask) to attack a face biometric system. The number of baseline anti-spoofing systems whose source code are publicly available for comparison and reproducible research is even more limited. Following the 1st competition on countermeasures to 2D spoofing attacks consisting of printed photographs, Tabula Rasa organized the 2nd competition [49] for comparison of the performance of spoofing countermeasures for a more diverse set of attacks. This competition was the part of the 6th International Conference of Biometrics (ICB 2013, 4th - 7th June, 2013, Madrid, Spain). In the field of face recognition, a large number of spoofing attacks are emerging. Simply, the number of anti-spoofing systems is however limited. The aim of the 2nd competition on countermeasures to 2D face spoofing attacks is to challenge researchers to create spoofing countermeasures effectively detecting a variety of face spoofing attacks. In first competition, face anti-spoofing algorithms were evaluated on Print-Attack database. After Print-Attack database [22], Replay-Attack database [28] is introduced with two additional types of attacks (photo attacks and video playbacks). In 2nd competition, face anti-spoofing algorithms were evaluated on Replay Attack database. Eight different teams from the universities around the world participated in the contest. Performance comparison between the teams has been presented in Table 5.

Teams: Eight teams those had participated in competition are –

i) Center for Biometrics and Security Research & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences (CASIA)

ii) Fraunhofer Institute for Computer Graphics Research (IGD)

iii) Joint team from Idiap Research Institute, University of Uoulu and University of Campinas (MaskDown)

iv) The LNM Institute Of Information Technology, Jaipur (LNMIIT)

v) Tampere University of Technology (MUVIS)

vi) University of Cagliari (PRA Lab)

vii) Universidad Autonoma de Madrid (ATVS)

viii) Joint team from University of Campinas and Universidad Federal de Minas Gerais (Unicamp)

Database and Protocol used:

Publicly available REPLAY-ATTACK database [28] was used to evaluate the performance of the algorithms presented by eight participating teams in 2nd competition on the face spoofing attack. In competition, initially, training and development data sets were given to contestants. The training set contains 60 real accesses, 60 printed attacks, 120 photo attacks and 120 video attacks. Development set also contains 60 real accesses, 60 printed attacks, 120 photo attacks and 120 video attacks. All the teams were given a couple of months to train and develop their classification system. The test data set contains 80 real access, 80 printed attacks, 160 photo attacks and 160 video attacks. All videos in the test set contained 100 frames with a random starting frame. To evaluate the performance of the algorithms properly, files of the test data set are anonymized to conceal the type of the video (real/attack). Organizers had asked all the teams to provide scores of the development set and test set by the specified deadline (March 15th, 2013). To compute the performance measure of the spoofing detection systems, at first a threshold at equal error rate (EER) on the development set scores is computed. Then, on the test set scores using the same threshold, they computed half total error rate (HTER), which combines the false rejection ratio (FRR) and the false acceptance ratio (FAR) with 0.5 weight. Spoof detection accuracy is good if anyone in FAR, FRR and HTER is close to 0%.

Methods:

1) CASIA

This team proposes a feature-level fusion of motion and texture characteristics. The motion feature was implemented using Gunnar Farneback's algorithm [73] to extract dense optical flows between two frames, with an interval of five frames and the texture features were extracted using multi-scale Local Binary Patterns (LBP) as in [11] to analyze the quality degradation of the attack samples. The extracted motion and texture features are concatenated into the final feature vector and finally, the features are fed into a linear Support Vector Machine (SVM).

2) IGD

This team implemented an algorithm based on Eulerian magnification [74] to magnify the small changes of color and movement which appear on the face due to the blood flow. Eulerian magnification uses two main approaches for spatial decomposition: Laplacian and Gaussian pyramids. The outputs obtained from Gaussian and Laplacian filters were processed with PCA for dimensionality reduction and produce two final feature vectors. The two feature vectors are then fed into separate AdaBoost classifiers. The obtained scores for the two approaches are then normalized and combined using weighted sum fusion rule.

3) MaskDown

This team introduced a composite system by joining 4 different categories of counter-measures together, each of which may be effective against a single type of attack. First two approaches based on Uniform Local Binary Pattern (LBP) [12] and Grey-Level Co-occurrence Matrix (GLCM) [24] was used to extract textural features. These features are classified using Linear Discriminant Analysis (LDA). The next approach used LBP-TOP [75] to extract textural patterns in a dynamic way. The fourth counter-measure is motion based and exploits the high correlation between the background and the face region movements in the case of spoofing attacks. The textural features are computed for each frame separately on normalized face bounding box. The score for the full videos is obtained by averaging the scores for each frame. Finally, the scores for each video are fused using Linear Logistic Regression (LLR).

4) LNMIIT

This team utilizes three types of intuitive visual features for their algorithm. Firstly, LBP was used to capture textural information. Secondly, as a complementary non-rigid motion analysis approach, the team extracts face background consistency feature [76]. Gaussian Mixture Model (GMM) [77] is used for background modeling. The third type of features combine motion and texture analysis and consists of 2D FFT on the GMM modeled background. These three types of feature vectors are computed on a per-frame basis and then averaged over the full video. The three features vectors are ten combined together into a single feature vector. At the end, the Hidden Markov Support Vector Machines (SVM^{hmm}) [78].

5) MUVIS

This team used two types of texture features: LBP and Gabor. For the LBP features, rotation-invariant uniform LBP histogram [12] is computed individually for every frame of the video V, where, the number of neighboring pixels P = 16 and the radius R = 2. The Gabor features require computation of Gabor wavelet transform in 4 scales and 6 orientations. The feature vector on per-frame basis is constructed using mean and standard deviation of the magnitude of the transform coefficients [79]. The result of the concatenation of the two feature vector is fed in Partial Least Square regression [80] for final classification.

6) PRA Lab

This team explored several different types of visual features, e.g. color features, edges, textures etc. in their algorithm [53]. The features are extracted from each frame and they are used to train several SVMs for each feature separately.

7) ATVS

This team used 25 Image Quality Measures (IQM) which provide a quantitative score that describes the level of distortion of the input image. They are classified using LDA classifier. The system works on a frame – by – frame basis, and the final video score is produced as the average of all the frame scores in the sequence.

8) Unicamp

This team used visual rhythm technique [21]. This team extracts a set of features using the Gray Level Co-occurrence Matrices (GLCM) [24] from the image representing the visual rhythm. For classification stage SVM with radial basis function kernel was used.

Table 5. Techniques used by the teams and performance comparison between the teams who participated in the ICB 2013 competition on counter measure to 2D facial spoofing attacks.

Method	Techniques used	Development set			Test set		
		FAR	FRR	HTER	FAR	FRR	HTER
CASIA	Motion Analysis, Texture Analysis	0.00	0.00	0.00	0.00	0.00	0.00
IGD	Liveness Detection	5.00	8.33	6.67	17.00	1.25	9.13
MaskDown	Motion Analysis, Texture Analysis	1.00	0.00	0.50	0.00	5.00	2.50
LNMIIT	Motion Analysis, Texture Analysis	0.00	0.00	0.00	0.00	0.00	0.00
MUVIS	Texture Analysis	0.00	0.00	0.00	0.00	2.50	1.25
PRA Lab	Texture Analysis	0.00	0.00	0.00	0.00	2.50	1.25
ATVS	Texture Analysis	1.67	0.00	0.83	2.75	21.25	12.00
Unicamp	Texture Analysis	13.00	6.67	9.83	12.50	18.75	15.62

7 Different Face Spoofing Databases

Currently, there are several face spoofing databases are available. They are IDIAP Replay-Attack database [28], IDIAP Print-Attack database [22], CASIA Face Anti-Spoofing database [9], NUAA Photograph Imposter database [14], 3D Mask Attack database [46] and MSU MFSD [46].

7.1 Replay-Attack database

Replay-Attack database [28] (Figure 15) is a face spoofing database, introduced by IDIAP Research Institute, Switzerland. This database contains 1300 video clips of photo and video attack attempts to 50 clients. Replay video attack and real access attempt videos are captured for at least 9 seconds at the frame rate of 25 Hz by webcam of a Macbook laptop, which produces colour videos with a resolution of 320×240 pixels, using the QuickTime framework and saved as ".mov" files. Real access and replay attack videos are captured under controlled and adverse lighting condition. Attacks were performed by high resolution photos and videos of each client, which were taken under the same conditions as in their authentication sessions. A Canon PowerShot *S*×150 IS camera is used for this purpose, which records 12.1 Mpixel photographs and 720p high-definition video clips. There are two subsets of attacks. The first sets of videos are generated by a stand, holding the client's biometry and the second set of videos are generated by holding the device with their own hand.

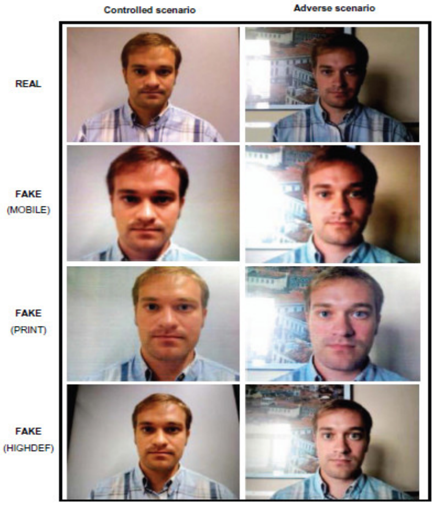

Figure 15. Typical examples of real and fake (mobile, print and highdef) face images of Replay-Attack database. Images were extracted from videos acquired in the two considered scenarios: controlled and adverse [56].

7.2 Print-Attack database

Print-Attack database [22] is a face spoofing database, introduced by IDIAP Research Institute, Switzerland. This database contains 200 video clips of printed photo attack and 200 real-access attempt videos of 50 clients. Printed photo attack and real access attempt videos are captured for at least 9 seconds at the frame rate of 25 Hz by webcam of a Macbook laptop, which produces colour videos with a resolution of 320×240 pixels using the QuickTime framework and saved as ".mov" files. Real access and print attack data are collected under both controlled and adverse lighting conditions. Attacks are performed by high resolution printed photographs of the clients which are taken under the same conditions as in their authentication sessions. There are two subsets of attacks. The first set of videos is generated by a stand, holding the client's printed photo and the second set of videos are generated by holding the picture with their own hand. Sample face images taken in adverse and controlled environments from Print Attack database are shown in Figure 16.

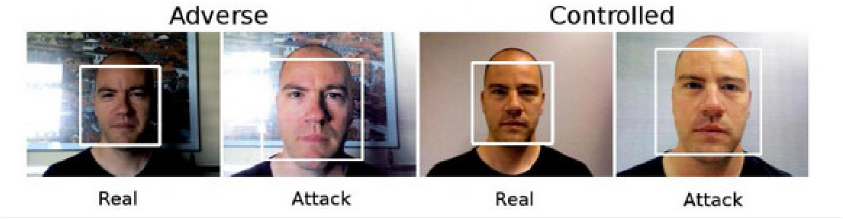

Figure 16. Sample images from Print Attack database [22] are shown.

7.3 CASIA Face Anti-Spoofing Database

CASIA Face Anti-Spoofing Database [9] is introduced by Center for Biometric and Security Research. This database contains 600 video clips (3 genuine and 9 fake videos of each person) of 50 persons. Fake faces are produced by high quality records of the genuine faces. Three imaging qualities (low, normal, and high) are considered here. Three fake face attacks (warped photo attack, cut photo attack and video attack) are implemented in this database. One complete video set for an individual from the CASIA Face Anti-Spoofing database is shown in Figure 17.

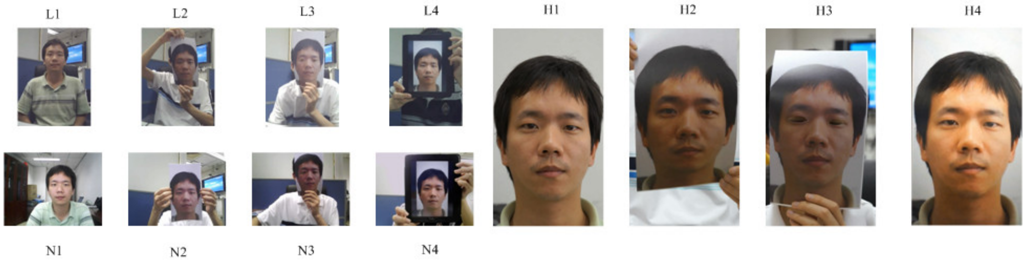

Figure 17. One complete video set for an individual subject in CASIA-FAS DB.

7.4 NUAA PI Database

NUAA PI database [14] contains real faces and photographs of corresponding 15 persons. The database is collected in a different place and in different illumination conditions having three sessions with about two weeks interval between any two sessions, using several unspecified webcams. 500 images of each subject are captured with frontal view, neutral expression and no apparent movements by webcam with the frame rate of 20 fps. To collect the samples, at first, high definition photos captured using a usual Canon camera where face area occupies 2/3 of the whole face area. Then photos are developed in two ways – (a) traditional print method is used on a photographic paper with size 6.8 cm×10.2 cm and 8.9 cm×12.7 cm and (b) print each photo on a 70g A4 paper using a color HP printer. The real faces and the corresponding photographs from NUAA PI database are shown in Figure 18 and Figure 19.

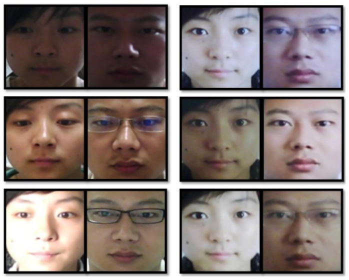

Figure 18. Illustration of the samples from the database [14]. In each column (from top to bottom) samples are respectively from session 1, session 2 and session 3. In each row, the left pair are from a live human and the right from a photo. Note that it contains various appearance changes commonly encountered by a face recognition system (e.g., sex, illumination, with/without glasses). All original images in the database are color pictures with the same definition of 640 x 480 pixels.

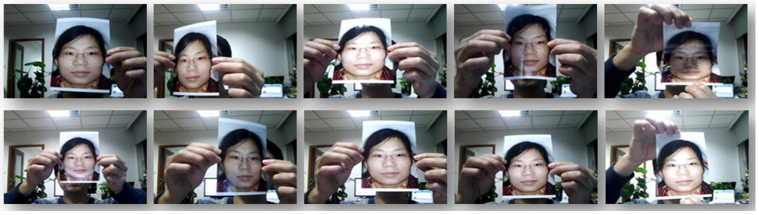

Figure 19. Illustration of different photo-attacks (from left to right) [14]: (1) move the photo horizontally, vertically, back and front; (2) rotate the photo in depth along the vertical axis; (3) the same as (2) but along the horizontal axis; (4) bend the photo inward and outward along the vertical axis; (5) the same as (4) but along the horizontal axis.

7.5 3D Mask Attack Database

3D Mask Attack Database (3DMAD) [46] is a face spoofing database, introduced by IDIAP Research institute, Switzerland. This database contains total 76500 frames of 17 persons for both real access and spoof attacks, which are recorded by Kinect in controlled lighting condition with neutral facial expression and frontal view. Samples are collected in three sessions – first two sessions capture real access samples and third session capture 3D mask attacks. A number of videos per subject is 5 and number of frames per video of a subject is 300. Description of each frame is given as follows.

- A depth image (640 X 480 pixels -1 X 11 bits)
- The corresponding RGB image (640 X480 pixels – 3 X8 bits)
- Manually annotated eye positions (with respect to the RGB image)

In each video, the eye positions are manually labeled in every 1st, 61st, 121st, 181st, 241st and 300th linearly interpolated for the rest.

7.6 MSU MFSD Database

MSU MFSD [40] is a face spoofing database of Pattern Recognition and Image Processing (PRIP) Lab at the Michigan State University, US. The database contains 280 video clips of photo and video attempts of 35 clients. Built-in camera in MacBook Air 13” (640×480) laptop and front facing the camera in the Google Nexus5 Android phone (720×480) are used to collect this database. All videos are generated by real access or printed photo or a video recording (at least for 9 seconds) of the client. Printed photo for attacks is generated with state of the art color printer on larger sized paper. The printed photos in this database have much better quality than the printed photos in the IDIAP and CASIA databases. Videos captured by Macbook Air 13” laptop are “.mov” files and videos captured by Google Nexus Android phone are “.mp4” files. Example faces from MSU MFSD database are shown in Figure 20.

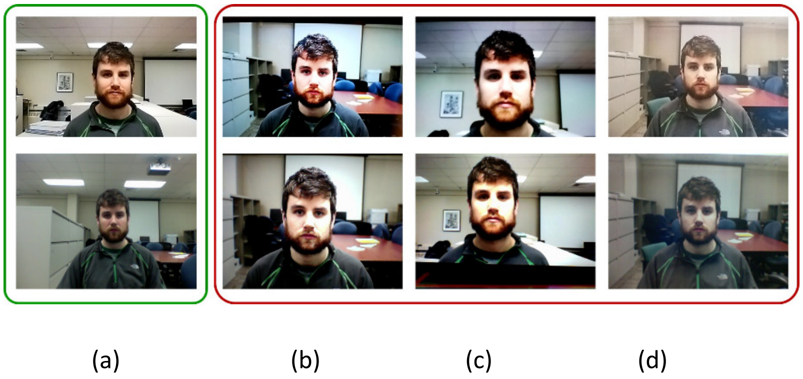

Figure 20. Example images of genuine and spoof faces of one of the subjects in the MSU MFSD database [40] captured using Google Nexus 5 smart phone camera (top row) and MacBook Air 13" laptop camera (bottom row). (a) Genuine faces; (b) Spoof faces generated by iPad for video replay attack; (c) Spoof faces generated by iPhone for video replay attack; (d) Spoof faces generated for printed photo attack.

8 Comparisons between Different Face Spoof Detection Techniques

This section presents a comparative study of some well-studied face spoof detection techniques discussed in Section 5.1. The objective of this comparative study is to show the most relevant characteristics of these techniques in order to get an inside into by looking at a glance at the table and it also shows an overall aspects of the different approaches studied so far in the field of face spoof detection. Table 6 demonstrates the comparisons among different face spoof detection techniques with some common attributes. Their performances are given in the table in terms of false reject rate (FRR), false accept rate (FAR), equal error rate (EER), half total error rate (HTER) and area under curve (AUC), and they are self-explanatory.

Table 6. Comparisons between different face spoof detection techniques.

Name of Method	Objective	Used Technique	Classifier	Database	Performance
Micro-Texture Analysis [11]	Detect face print attack	Multi-scale LBP	SVM classifier with RBF kernel	NUAA PI database	AUC=0.99% Classification accuracy=98.0% FAR =0.6% and FRR =4.4%
Video-Based Face Spoofing Detection through Visual Rhythm Analysis [20]	Detect video based face spoofing attack	Visual Rhythm + Gray Level Co-occurrence Matrices (GLCM)	SVM classifier with Linear and RBF kernel + Partial Least Squares(PLS) Regression method	NUAA PI database and Print Attack database	AUC=100%
Motion Magnification [26]	Detect face print attack and replay attack	i) Motion Magnification with Multi- scale LBP and ii) Motion Magnification with HOOF	i) SVM classifier with RBF kernel ii) LDA	Print Attack database and Replay Attack database	HTER=1.25% (for print attack) and HTER=6.62%(for replay attack) HTER=0% (for print attack) and HTER=1.25%(for replay attack)
Image Distortion Analysis [33]	Detect print photo attack and replay video attack	Image Distortion Analysis based algorithm using four features(specular reflection feature, blurriness feature, chromatic moment feature, and color diversity feature)	Ensemble classifier (multiple SVM)	Replay Attack database, CASIA-FASD and MSU MFSD	Intra-database Testing scenario – HTER=7.41% ,TPR=92.2% @ FAR=0.1, and TPR=87.9% @ FAR=0.01(for Replay Attack database), EER = 5.82%,TPR=94.7% @ FAR=0.1 and TPR =82.9% @ FAR=0.01 (for MSU MFSD database using 55 subjects) EER=8.58%,TPR=92.8% @ FAR=0.1 and TPR =64.0% @ FAR=0.01 (for MSU MFSD database using 35 subjects) EER=13.3%(30frms),TPR=86.7% @ FAR=0.1 and TPR =50% @

					FAR=0.01 (for CASIA database –H protocol) EER=12.9%(75frms),TPR=86.7% @ FAR=0.1 and TPR =59.7% @ FAR=0.01 (for CASIA database –H protocol) Cross-database Testing scenario –For replay attack – TPR=90.5% @ FAR=0.01 (MSU vs. Idiap database) For printed attack – TPR=31.2% @ FAR=0.01 (MSU vs. Idiap database)
Live Face Video vs. Spoof Face Video: Use of Moire Patterns to Detect Replay Video Attacks [44]	Detect Spoof face video attack	Multi-scale LBP +Densely Sampled SIFT (DSIFT)	SVM classifier with a RBF kernel	Replay Attack database, CASIA FASD database, and RAFS based on the MSU-FSD database	Intra-database Testing scenario –HTER = 18.0% (on Idiap), 49.0% (on CASIA), and 11.4% (on RAFS) Cross-database Testing scenario – 3.3% (on Idiap database), 0.0% (on CASIA FASD), and 11.3% (on RAFS) Classification Accuracy – 88.0% (on Idiap database), 67.3% (on CASIA FASD), and 85.5% (on RAFS database)
Texture and Local Shape Analysis [5]	Detect photo attack	LBP + Gabor Wavelet + HOG	Linear SVM	NUAA Photograph Imposter Database, Yale Recaptured Database, Print-Attack Database	EER= 1.1% and AUC=0.999 on NUAA PI DB Accuracy=100%, Positive rate= 100% and False Positive Rate = 0.0% on Yale Recaptured DB FAR= 1.67, FRR=1.67 on development set of Print-Attack DB FAR=FRR=HTER=0.0% on test set of Print-Attack DB
General Image Quality Assessment [56]	Detect photo attack	14 Image Quality Measure(9 Pixel Difference measure, 3 correlation based measure, 1	Linear Discriminant Analysis (LDA)	Replay-Attack DB, CASIA FAS-DB	HTER=2.8 for hand-based mobile photo, HTER=9.3 for hand-based print photo and HTER=13.1 for hand-based highdef photo on Replay-Attack DB

		edge based and 1 corner based measure)			HTER=3.5 for fixed-support mobile photo, HTER=8.4 for fixed-support print photo and HTER=9.1 for fixed-support highdef photo on Replay-Attack DB HTER= 25.0 for low-resolution warped photo, HTER =23.3 for low-resolution cut photo and HTER=21.7 for low-resolution video on CASIA-FAS DB. HTER= 23.3 for normal-resolution warped photo, HTER =16.7 for normal-resolution cut photo and HTER=23.3 for normal-resolution video on CASIA-FAS DB. HTER= 10.0 for high-resolution warped photo, HTER =11.7 for high-resolution cut photo and HTER=6.7 for normal-resolution video on CASIA-FAS DB.
Face Liveness Detection for Combating the Spoofing Attack in Face Recognition [57]	Detect photo attack and replay video attack	Dynamic High Frequency Descriptor (DHFD)	Binary Classifier	Author's private database	Precision=95% when resolution is 500×500 and 250×250 Precision = 85.71% when the resolution is 1000×1000 (precision is the fraction of retrieved positive samples that are relevant)
Motion Based Counter-Measure to Photo Attacks in Face Recognition [60]	Detect Photo Attack	Optical Flow Correlation (OFC)	Binary Classifier	Photo-Attack Database	EER=0.36% and HTER=1.52% on development set when Full Background is considered EER=0.06% and HTER=1.76% on test set when the background size is reduced to 50%.

The techniques which are listed in Table 6 have some merits and also have some demerits. The method which is discussed in [11] is a texture analysis based face spoof detection technique. This approach uses the texture patterns that provide detectable information between the texture of real and fake faces. This technique is very simple to implement, implementation cost is low, no user collaboration is needed here and it produces a good result in known scenarios. But this technique needs data that covers all possible attacks. As different type of attacks generates different texture patterns, feature space used for face spoof detection should be large. Low-quality video or bad quality image does not give good result in this case. The method exploited in [20] is also a texture analysis based method. This method uses the visual rhythms that can be interpreted as texture maps and it can be summarized using simple texture descriptors, such as the gray level co-occurrence matrices. This method responds to input fast in small and medium computational systems. The reduced dimensionality of the method allows the use of this method in large video databases. But the low quality video is a big issue for this method. The work presented in [26] introduces two different algorithms for face spoof detection, among these two algorithms, one is based on texture analysis (using Multi-scale LBP) and another is based on motion estimation (using HOOF). Before applying the texture analysis or motion estimation technique, Eulerian motion magnification approach is used to enhance the facial expressions of a subject in a captured video. This motion magnification improves the performance of Multi-scale LBP texture features and HOOF descriptor. The approach outperforms existing approaches in terms of accuracy as well as computational time. But this approach faces some problems when a short duration video of a person is displayed or simply shaking the photograph in front of the camera. The work discussed in [33] is an image distortion (image quality) analysis based approach. This is an efficient and robust algorithm. This algorithm has good generalization capability and low computational complexity. But an ensemble classifier (different classifiers needed for different spoof attack) needed for this technique. The algorithm reported in [44] is used to detect the spoof face video with the help of moiré pattern which appears when a video is recaptured or photo replays on a screen. In this technique without face detection operation spoof face detection is possible, because moiré patterns exist in the whole video frame containing the face. Due to this reason, this approach is useful in a challenging environment where non-frontal faces present in front of the camera. The characteristics of moiré patterns are represented using Multi-scale LBP and DSIFT in this technique. This technique is very simple to implement, the computational cost is low and no user collaboration is needed. But low-quality video does not give a good result. The approach proposed in [5] is robust, computationally fast and does not require user cooperations. The texture features that are used for face spoof detection can also be used for face recognition. This provides a unique feature space for coupling spoof detection and face recognition. In this case, the feature space used for face spoof detection should be large, because face images captured from printed photographs may visually look very similar to the face images captured from live faces and all these images are found largely overlapping in the original input space. The approach reported in [56] is very simple, non-intrusive and user-friendly. The speed of this approach is high and complexity is low, which is very desirable in a practical protection system. This anti-spoofing approach does not require any preprocessing steps and relies only on general image quality measures, therefore, it may be used for different biometrics modalities (e.g., iris, fingerprint or palmprint). The proposed approach degrades when the quality of the spoofing attempts increases. The face-spoofing method introduced in [57] is a liveness detection based method. This method is non-intrusive and easy to implement. The performance of this approach degrades when high resolution screens are used to attack the system. The approach discussed in [60] is a motion based countermeasure to face spoofing. This method requires the estimation of Optical Flow (OF) fields, which can dramatically improve the accuracy of spoofing detectors. The performance of this approach is dependent on the background size.

9 Conclusion

At present face biometric systems are not adequately protected to spoof attacks and the most common sources of spoof identity attacks are photographs and videos. Different researchers proposed different techniques to protect the face biometric system from spoofing attacks. The face spoof detection techniques which are discussed in this paper are using texture pattern analysis, motion analysis, life sign analysis and image quality analysis. However, none of them are not free from sensitive constraints and limitations. Hence, advancement is needed in this field to provide a robust and computationally efficient solution to improve the use of face biometric. In this paper, we present a detail survey of different face spoof detection techniques, their implementation details, performance and limitations which are beneficial for researchers in this field. We also summarize some publicly available face spoofing databases and conclude our discussion by comparing some well-studied techniques.

ACKNOWLEDGEMENTS

This study presented in this paper is partially supported by Sir Visvesvaraya Young Faculty Research Fellowship grant from Media Lab Asia, Department of Electronics and Information Technology, Government of India.

REFERENCES

[1] W. Bao, H. Li, N. Li, and W. Jiang. A liveness detection method for face recognition based on optical flow field. In *2009 International Conference on Image Analysis and Signal Processing*, pages 233–236. IEEE, 2009.

[2] K. Kollreider, H. Fronthaler, and J. Bigun, “Non-intrusive liveness detection by face images," *Image and Vision Computing*, 27: 233–244, 2009.

[3] T. F. Pereira, A. Anjos, J. De Martino, and S. Marcel, “Can face anti-spoofing countermeasures work in a real world scenario?,” In Proc. ICB, pages 1–8, June 2013.

[4] M. De Marsico, M. Nappi, D. Riccio, and J. Dugelay, “Moving face spoofing detection via 3D projective invariants,” In *International Conference on Biometrics*, pages 73–78, 2012.

[5] J. Määttä, A. Hadid, and M. Pietikainen, “Face spoofing detection from single images using texture and local shape analysis,” *IET Biometrics*, 1(1):3–10, 2012.

[6] T. d. F. Pereira, A. Anjos, J. M. De Martino, and S. Marcel, “ LBP-TOP based countermeasure against facial spoofing attacks,” In *ACCV Workshop on Computer Vision With Local Binary Pattern Variants*, 2012.

[7] Peixoto, C. Michelassi, and A. Rocha, “Face Liveness Detection under Bad Illumination Conditions,” in IEEE Intl. Conference on Image Processing, Sep. 2011, pp. 3557–3560.

[8] W. R. Schwartz, A. Rocha, and H. Pedrini, “Face Spoofing Detection through Partial Least Squares and Low-Level Descriptors,” in Intl. Joint Conference on Biometrics, Oct. 2011, pp.1–8

[9] Stauffer and W. Grimson, “Adaptive background mixture models for real-time tracking,” In Computer Vision and Pattern Recognition, 1999.

[10] G. Pan, L. Sun, Z. Wu, and Y. Wang, "Monocular Camera-based Face Liveness Detection by Combining Eye blink and Scene Context," Telecommunication Systems, vol. 47, pp. 215–225, 2011.

[11] J. M¨a¨atta, A. Hadid, and M. Pietik¨ainen, "Face spoofing detection from single images using micro-texture analysis," In Proc. IJCB, pages 1–7, 2011.

[12] T. Ojala, M. Pietik¨ainen, and T. M¨aenp¨a¨a, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, 24:971–987, July 2002.

[13] T. Ahonen, A. Hadid, and M. Pietik¨ainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, 28:2037–2041, December 2006.

[14] X. Tan, Y. Li, J. Liu, and L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model," In *Proceedings of the 11th European conference on Computer vision: Part VI*, ECCV'10, pages 504–517, Berlin, Heidelberg, 2010. Springer-Verlag.

[15] V. N. Vapnik, "*Statistical Learning Theory,"*Wiley-Interscience, 1998.

[16] B. S.ManjunathandW. Y.Ma., "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach.Intell.*, 18:837–842, August 1996.

[17] C.-C. Chang and C.-J. Lin., "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/˜cjlin/libsvm

[18] V. Ojansivu and J. Heikkil¨a., "Blur insensitive texture classification using local phase quantization," In *Proceedings of the3rd international conference on Image and Signal Processing*, ICISP '08, pages 236–243, Berlin, Heidelberg, 2008. Springer-Verlag.

[19] J. Yang, Z. Lei, S. Liao, and S. Li., "Face liveness detection with component dependent descriptor," In Proc. ICB, pages 1–6, June 2013.

[20] A. d. S. Pinto et al., "Video-based face spoofing detection through visual rhythm analysis," in 25th Conference on Graphics, Patterns and Images, 2012.

[21] M.-G. Chung, J. Lee, H. Kim, S. M.-H. Song, and W.-M. Kim, "Automatic Video Segmentation based on Spatio – Temporal Features," Korea Telecom, vol. 1, no. 4, pp. 4–14, 1999.

[22] A. Anjos and S. Marcel, "Counter-Measures to Photo Attacks in Face Recognition: A Public Database and A Baseline," in Intl. Joint Conference on Biometrics, 2011, pp. 1–7.

[23] H. Wold, S. Kotz and N. Johnson, "Partial Least Squares," in Encyclopedia of Statistical Sciences," New York: Wiley, 1985, vol. 6, pp. 581–591.

[24] R. Haralick, K. Shanmugam, and I. Dinstein, "Texture Features for Image Classification," IEEE Trans. on Systems, Man, and Cybernetics, vol. 3, no. 6, 1973.

[25] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Trans. on Communications*, vol. 43, pp. 2959–2965, 1995.

[26] S. Bharadwaj, T. I. Dhamecha, M. Vatsa, and R. Singh., "Computationally efficient face spoofing detection with motion magnification," In Proc. CVPR Workshops, pages 105–110, 2013.

[27] R. Chaudhry, A. Ravichandran, G. Hager, and R. Vidal, “Histograms of oriented optical flow and Binet-Cauchy kernels on nonlinear dynamical systems for the recognition of human actions,” In *Conference on Computer Vision and Pattern Recognition*, pages 1932–1939, 2009.

[28] I. Chingovska, A. Anjos, and S. Marcel, “On the effectiveness of local binary patterns in face anti-spoofing,” In *International Conference of the Biometrics Special Interest Group*, pages 1–7, 2012.

[29] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman, “Eulerian video magnification for revealing subtle changes in the world,” *ACM Transactions onGraphics*,31(4),2012.

[30] M. Shreve, S. Godavarthy, D. Goldgof, and S. Sarkar, “Macro and micro-expression spotting in long videos using spatiotemporal strain,” In *International Conference on Automatic Face Gesture Recognition and Workshops*, pages 51–56, 2011.

[31] Y. Altun, I. Tsochantaridis, and T. Hofmann, “Hidden markov support vector machines,” In International Conference on Machine Learning, 2003.

[32] M. G. Martini *et al.*, “Image quality assessment based on edge preservation,” *Signal Processing: Image Communication*, vol. 27, pp. 875–882, 2012.

[33] D. Wen, H. Han, and A. K. Jain, “Face Spoof Detection with Image Distortion Analysis,” *IEEE Transactions on Information Forensics and Security*, Vol. 10, No. 4, pp.746-761, April 2015.

[34] Pittsburgh Pattern Recognition (PittPatt), PittPatt Software Developer Kit (acquired by Google), http://www.pittpatt.com/.

[35] R. Tan and K. Ikeuchi, “Separating reflection components of textured surfaces using a single image,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 178–193, Feb. 2005.

[36] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas, “The blur effect: perception and estimation with a new no-reference perceptual blur metric,” in *Proc. SPIE: Human Vision and Electronic Imaging XII*, 2007.

[37] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, “A no-reference perceptual blur metric,” in *Proc. ICIP*, vol. 3, 2002, pp. 57–60.

[38] Y. Chen, Z. Li, M. Li, and W.-Y. Ma, “Automatic classification of photographs and graphics,” in *Proc. ICME*, 2006, pp. 973–976.

[39] G. Pan, L. Sun, Z. Wu, and S. Lao, “Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Web camera,” in IEEE Intl. Conference on Computer Vision, 2007, pp. 1–8.

[40] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, “A face anti spoofing database with diverse attacks,” in *Proc. ICB*, 2012, pp. 26–31.

[41] D. Lowe, “Object recognition from local scale-invariant features,” In Proc. ICCV, pages 1150–1157, 1999.

[42] N. Kose and J.-l. Dugelay, “Classification of captured and recaptured images to detect photograph spoofing,” in *Proc. ICIEV*, 2012, pp. 1027–1032.

[43] J. Galbally, S. Marcel, and J. Fierrez, “Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition,” *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 710–724, 2014.

[44] K. Patel, H. Han, and A. K. Jain, “Live Face Video vs. Spoof Face Video: Use of Moire Patterns to Detect Replay Video Attacks,” *ICB*, Phuket, Thailand, May 19-22, 2015.

[45] A. Amidror, “The Theory of the Moir´e Phenomenon Volume I: Periodic Layers,” 2nd ed. Springer, 2009.

[46] https://www.IDIAP.ch/dataset

[47] Dalal, N., Triggs, B., “Histograms of oriented gradients for human detection,” Int. Conf. on Computer Vision & Pattern Recognition, 2005, vol. 2, pp. 886–893

[48] Vedaldi, A., Zisserman, A., “Efficient additive kernels via explicit feature maps,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2010

[49] A. Chingovska, J. Yang, Z. Lei, D. Yi, S. Z. Li, O. Kähm, C. Glaser, N. Damer, A. Kuijper, A. Nouak, J. Komulainen, T. Pereira, A. Anjos, S. Gupta, S. Khandelwal, S. Bansal, A. Rai, T. Krishna, D. Goyal, M.-A. Waris, H. Zhang, I. Ahmad, S. Kiranyaz, M. Gabbouj, R. Tronci, M. Pili, N. Sirena, F. Roli, J. Galbally, J. Fierrez, A. Pinto, H. Pedrini, W. S. Schwartz, A. Rocha, and S. Marcel, “The 2nd competition on counter measures to 2d face spoofing attacks,” In *IAPR International Conference on Biometrics (ICB)*, 2013.

[50] S. de Jong., “SIMPLS: an alternative approach to partial least squares regression,” Chemometrics and Intelligent Laboratory Systems, 18(3):251–263, 1993.

[51] Viola, P.A., Jones, M.J., “Rapid object detection using a boosted cascade of simple features,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2001, pp. 511–518.

[52] Niu, Z., Shan, S., Yan, S., Chen, X., Gao,W., “2D cascaded adaboost for eye localization,” Proc. 18th Int. Conf. on Pattern Recognition, 2006.

[53] Chakka, M.M., Anjos, A., Marcel, S., et al., “Competition on counter measures to 2-d facial spoofing attacks,” Proc. IAPR IEEE Int. Joint Conf. on Biometrics (IJCB), Washington, DC, USA, 2011.

[54] Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., Lin, C.J., “LIBLINEAR: a library for large linear classification,” J. Mach. Learn. Res., 2008, 9, pp. 1871–1874.

[55] Vedaldi, A., Fulkerson, B., “VLFeat: an open and portable library of computer vision algorithms,” 2008, http://www.vlfeat.org/

[56] Javier Galbally, Sébastien Marcel, “Face Anti-Spoofing Based on General Image Quality Assessment,” Pattern Recognition (ICPR), 2014 22nd International Conference, pp. 1173-1178, 2014.

[57] Junyan Peng, Patrick P.K. Chan, “Face Liveness Detection For Combating The Spoofing Attack in Face Recognition,”2014 International Conference on Wavelet Analysis and Pattern Recognition, pp. 176-181, 2014.

[58] S. Yao *et al.*, “Contrast signal-to-noise ratio for image quality assessment,” in *Proc. Int. Conf. on Image Processing*, 2005, pp. 397–400.

[59] Kollreider, Klaus, Hartwig Fronthaler, and Josef Bigun."Verifying liveness by multiple experts in face biometrics," In Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, pp. 1-6. Ieee, 2008.

[60] André Anjos, Murali Mohan Chakka , Sébastien Marcel, “Motion-based counter-measures to photo attacks in face recognition”, IET Biometrics, vol. 3, pp. 147-158, 2014.

[61] C. Liu, Ed., “Beyond Pixels: Exploring New Representations and Applications for Motion Analysis,” Doctoral Thesis. Massachusetts Institute of Technology, May 2009.

[62] A. Bruhn, J. Weickert, and C. Schn¨orr, “Lucas/kanade meets horn/schunck: Combining local and global optic flow methods,” International Journal of Computer Vision, vol. 61, pp. 211–231, 2005.

[63] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, “High accuracy optical flow estimation based on a theory for warping,” in European Conference on Computer Vision (ECCV). Springer, 2004, pp. 25–36.

[64] R. Tronci, G. Giacinto, and F. Roli, “Dynamic score combination: A supervised and unsupervised score combination method,” In Machine Learning and Data Mining in Pattern Recognition, volume 5632 of Lecture Notes in Computer Science, pages 163-177. Springer Berlin / Heidelberg, 2009.

[65] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma., “RASL: Robust Alignment by Sparse and Low-rank Decomposition for Linearly Correlated Images,” In IEEE Conference on Computer Vision and Pattern Recognition, pages 763-770, 2010.

[66] D. Donoho, “De-noising by Soft-thresholding,” IEEE Trans. on Information Theory, 41(3):613-627, 1995.

[67] D. Donoho and I. Johnstone, “Ideal Spatial Adaptation via Wavelet Shrinkage,” Biometrika, 81(3):425-455, 1994.

[68] W. R. Schwartz, R. D. da Silva, L. S. Davis, and H. Pedrini, “A Novel Feature Descriptor Based on the Shearlet Transform,” In IEEE Intl. Conference on Image Processing, 2011.

[69] Q. Huynh-Thu and M. Ghanbari, “Scope of validity of PSNR in image/video quality assessment,” *Electronics Letters*, vol. 44, pp. 800–801, 2008.

[70] C. Chow and C.N.Liu, “Approximating Discrete Probability Distributions with Dependence Trees,” IEEE Trans. on Information Theory, 14:426-467, 1968.

[71] J. Pearl and S. Russell, “Handbook of Brain Theory and Neural Networks,” chapter Bayesian Networks. Cambridge, 2002.

[72] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA Data Mining Software: An Update,” SIGKDD Explorations, 11(1), 2009.

[73] G. Farneb¨ack, “Two-frame motion estimation based on polynomial expansion,” In SCIA, pages 363–370, 2003.

[74] Avcibas *et al.*, “Statistical evaluation of image quality measures,” *Journal of Electronic Imaging*, vol. 11, pp. 206–223, 2002.

[75] G. Zhao and M. Pietik¨ainen, “Dynamic texture recognition using local binary patterns with an application to facial expressions,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(6):915–928, 2007.

[76] Yan, Z. Zhang, Z. Lei, D. Yi, and S. Li., “Face liveness detection by exploring multiple scenic clues,” In 12th International Conference on Control, Automation, Robotics and Vision, 2012.