2D Face Liveness Detection: an Overview

Olga Kähm, Naser Damer

Competence Center Identification and Biometrics
Fraunhofer Institute for Computer Graphics Research IGD
Fraunhoferstraße 5
64283 Darmstadt, Germany
olga.kaehm@igd.fraunhofer.de
naser.damer@igd.fraunhofer.de

Abstract: Face recognition based on 2D images is a widely used biometric approach. This is mainly due to the simplicity and high usability of this approach. Nonetheless, those solutions are vulnerable to spoof attacks made by non-real faces. In order to identify malicious attacks on such biometric systems, 2D face liveness detection approaches are developed. In this work, face liveness detection approaches are categorized based on the type of liveness indicator used. This categorization helps understanding different spoof attacks scenarios and their relation to the developed solutions. A review of the latest works dealing with face liveness detection works is presented. A discussion is made to link the state of the art solutions with the presented categorization along with the available and possible future datasets. All that aim to provide a clear path for the future development of innovative face liveness detection solutions.

1 Introduction

Biometrics tries to identify or to verify individuals based on their behavior or physical characteristics. Face recognition is a widely used biometric technique with promising solutions and application fields. Face recognition approaches can be classified by the type of used acquisition. 3D face recognition is known for achieving high recognition rate and for being highly secure, especially against spoof attacks [ZSBF08]. 3D face recognition, however, requires the use of highly complicated and expensive devices. It also requires a large computation effort and, in most cases, the collaboration of the user.

2D face recognition approaches provide a low cost and highly usable recognition system. This is due to the simple hardware needed for acquisition (e.g a simple optical camera), the low computation effort needed for the algorithms and the non-required collaboration of the user. This makes 2D face recognition a perfect solution for ubiquitous biometrics and biometric solutions for embedded and mobile devices.

Being easily spoofed by face images is one of the main problems in the field of 2D face recognition solutions. In an effort to prevent spoofing attacks, many researchers presented different approaches to classify face images into real faces and fake or printed face images. Those approaches are usually referred to as face liveness detection.

Although liveness detection may refer to the detection of the faces properties of being alive (not dead), in this work, liveness detection refers to identify real faces from fake (non-real) faces unless otherwise mentioned.

In the next section, an approach categorization scheme based on liveness indicator types is introduced. In section 3, a review of the most interesting face liveness detection works is presented. Then, a discussion is built to link liveness indicator types, state of the art approaches, datasets, and future work. Finally, a conclusion is drawn.

2 Methodologies and Categorization

Solutions such as liveness detection can be discussed from different points of view. Those solutions can be categorized by the machine learning technique they use, their application fields, or the information (features) used. In this work, different types of information extracted from face data are discussed. That information will be referred to as liveness indicators. That information affects the performance of the solutions under different scenarios and application fields and is very important to avoid different types of spoofing.

Representing liveness detection with respect to the type of information used will help later identify different spoofing attacks and therefore different liveness detection algorithms to be developed, as well as the specification of future datasets that should be gathered to develop such solutions.

Here, face liveness detection approaches are separated into three main categories based on the type of liveness indicator used.

- Motion analysis - based on the clue, that planar objects (2D) move differently from real faces (3D objects). Those approaches are usually associated with optical flow calculations between different frames in a video sequence. Different patterns of optical flow fields are assumed to represent the differences between the movement of 3D faces and 2D faces. Here, real faces are assumed to have depth information (3D) and fake faces are assumed to be planar (2D).

- Texture analysis - based on the assumption that printed faces contained detectable texture patterns. Here, texture features are extracted from face images (single images or sequences) under the assumption that fake faces are printed, and the printing process or the material (paper) printed on produces certain texture patterns that do not exist in real faces.

- Life sign detection - tries to analyze signs of life from user images (eye blinking, lips movements). The developed algorithms under this approach focus on the movement of a certain identified part of the face. Challenge response approaches, where the user interaction is required, will also be dealt with as a sign of life.

3 Literature

In this section, some of the most interesting liveness detection approaches are presented.

One of those methods was introduced by Jee et al. [kJuJhY06]. The proposed algorithm for liveness detection is based on the analysis of the eyes movement. The basic assumption is that because of blinking and uncontrolled movements of the pupils in human eyes, there should be big shape variations. This should be suitable for optical flow calculation. This algorithm is further presented in the following. First, the center points of both eyes in the input image are detected. This step must be precise, because it is the initial step and it has much effect on the performance [kJuJhY06]. The authors used the fact that the intensity of the eye region is respectively lower. In order to find the candidates, a Gaussian filter is applied and then the local minimas of the intensities are found using a gradient descent algorithm. In the next step the candidate regions of interest are classified using Viola-Jones AdaBoost [VJ01] classifier and as a result, invalid eyes are removed.

After the centers of both eyes are found, a normalization of the face region around the eyes is required. Found regions are normalized to one size and a high pass filter (SQI-self quotient image [WLW04]) is applied. The results are 20x20 pixels images based on the centers of eyes. Then, the regions are binarized using a threshold obtained from the mean pixel values of the eye region. Eye regions from real faces have bigger variations in shape than regions obtained from fake faces.

Next step is the calculation of liveness score, those scores were calculated using hamming distances. In their experiment authors compared 5 left and 5 right eyes. Liveness score (hamming distance) between two frames is the number of different pixels between both regions. After the calculation of 10 liveness scores of both eyes, the average of scores is taken. If this average liveness score is bigger than a threshold, then the input image is a live face, otherwise it is a fake.

The authors captured 100 live faces and 100 fake faces of 10 persons. For the fake faces they used a photo printer and the size of the photographs was 120*100 mm. Results of the experiment are shown in Table 1. It can be concluded that live faces have a bigger distance and therefore larger variation and score than fake faces. The best results were achieved by setting the threshold to 21. In this case the FAR is 0.01 and FRR is 0.08.

TABLE I
HAMMING DISTANCE OF EYE REGIONS

	Hamming distance		
	Mean	Min	Max
Live face	30	18	47
Fake face	17	10	22

Table 1: Obtained hammingdistances [kJuJhY06]

Another algorithm was introduced by Bao et al. [BLLJ09]. The method proposed in this work uses optical flow. It analyzes the differences and properties of optical flow generated from 3D objects and 2D planes.

As seen in the Figure 1 the motion as it appears in images is a combination of four basic

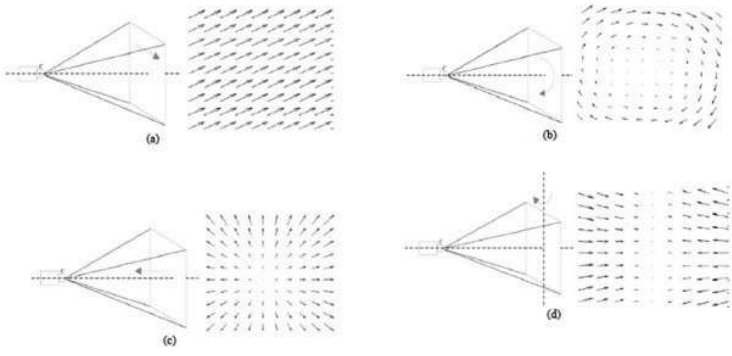

Figure 1: Four basic types of optical flow fields as presented by Bao et. al. [BLLJ09] a) Translation at constant distance from the observer b) Rotation at constant distance about the view axis c) Moving forward or backward d) Swing or rotation of a planar object perpendicular to the view axis

movement types: Translation, rotation, moving and swing.

The authors found that the first three basic types generate quite similar optical flow fields for both 2D and for 3D images. The fourth type creates more differences. Their approach is based on this idea. The optical flow field for 2D objects can be represented as a projection transformation. The optical flow allows to deduce the reference field, thus allowing to determine whether the test region is planar [BLLJ09]. For this, the difference between optical flow fields is calculated. To decide whether a face is a real face or not, this difference is thresholded. The Experiment was done on three groups of sample data. The first group contained 100 printed face pictures that were randomly rotated and translated, the second were 100 pictures from group 1 that were folded and curled before the test, the third group included faces of real people (10 people, each 10 times) doing gestures like swinging, shaking, etc.

The time for the experiment was 10 seconds. The camera had sampling rate of 30 frames per second. The system made calculation every 10 frames. Figure 2 shows examples of each group ((a)-group1, (b)-group 2 and (c)-group3) as well as the results obtained.

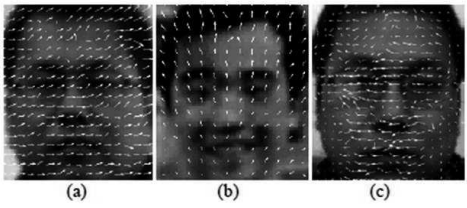

TABLE I
EXPERIMENTAL RESULTS

Group	T			
	0.2	0.4	0.6	0.8
1st	0.54	0.83	0.86	0.92
2nd	0.45	0.80	0.85	0.89
3rd	1.00	1.00	0.94	0.86

Figure 2: Examples (a)-group1, (b)-group 2 and (c)-group3 and Results: the ratio of successful detection in each experiment (0 to 1) [BLLJ09]

As shown in Figure 2, the greater the threshold (T) is, the higher the ratio of successful detection. But at a certain point the ratio will drop [BLLJ09], it must be noted that the authors did not mentioned any false acceptance rates. Another disadvantage is, that because the method is based on the precise calculation of the optical flow field, illumination changes will have a negative impact on the results. This method will also fail for 3-dimensional objects, because it is based on the assumption, that a fake face is a planar. Therefore, authors give advice to use this algorithm with other liveness detection methods [BLLJ09]. Another technique for liveness detection is introduced by Kollreider et al. [KFB]. It analyzes the trajectories of single parts of a live face. Obtained information can be used to decide whether a printed image was used. This approach uses a model-based Gabor decomposition and SVM for face parts detection [KFB]. To determine a liveness score, the authors combined face parts detection and an estimation of the optical flow of lines. This estimation is able to differentiate between motion of points and motion of lines. A detailed review of existing techniques for optical flow is given in [BFB94].

The idea of this method is based on the assumption that a 3D face generates a 2D motion which is higher at central face regions than at the outer face regions (e.g ears). Therefore, parts closer to the camera move differently from parts which are on a greater distance. In contrast to that, a photograph generates a constant motion on different face regions. With the knowledge of the face parts positions and their velocity, it is possible to compare how fast they are in relation to each other. This information is used to discriminate a live face against a photograph.

The authors proposed algorithms for the computing and implementation of the optical flow of lines (OFL). For this, they use the main Gabor filters that are linear filters for edge detection. The authors also introduce two approaches for the face parts detection: one is based on optical flow pattern matching and model-based Gabor feature classification. The second method extracts Gabor features in a non-uniform retinotopic grid and classifies them with trained SVM experts.

The used database contained 100 videos of Head Rotation Shot-subset (DVD002 media) of the XM2VTS database. All data was downsized to 300x240 pixels. Videos were cut (3 to 5 frames) and used for live and non-live sequences. The last frame from each person was taken and was translated horizontally and vertically to get 2 non-live sequences per person. Therefore 200 live and 200 non live sequences were analyzed. Most of the live sequences achieved a score of 0.75 out of maximum 1, whereas the non-live pictures achieved less than 0.5. It was also observed that glasses and moustaches lowered the score, because they were near to the camera. The authors mentioned that by using sequences containing only horizontal movements the system is error free. By considering a liveness score greater than 0.5 as alive, the system separates 400 test sequences with error rate of 0.75 % [KFB].

Kollreider et. al. introduced in [KFFB] an approach for face detection based on face-landmarks. One part of their work focuses on liveness detection using the detected face-landmarks. The method analyzes lip movements classifications (used SVM to classify lip dynamics) and lip-reading for the purpose of liveness detection in a dialogue scenario. The authors proposed to locate the mouth region and extract OFL in real-time. They presented results on the XM2VTS database simulating required scenarios (persons were recorded speaking digits from 0 to 9). The goal was to recognize the digits by lip-motion

only. As a database they used 100 short videos for every digit. The dataset was divided into 60 videos for training (SVM classifier, cross validation) and 40 for testing. During experiments features vectors are extracted from mouth regions and given to a 10-class SVM (for each of digit videos). As a result of the experiment, confusion matrices are obtained. Successful recognition rate out of 100 individuals is 0.73 (73%). The authors proposed it as indication of liveness. Face/mouth detection were also analyzed. Further results and details for the algorithms of this approach can be found in [KFFB].

In the context of the competition on counter measures to 2D face spoofing attacks [CAM+11], six teams have participated. A database was collected and made publicly available (PRINT-ATTACK) [1] [AM11]. The database contains 200 videos of real accesses and 200 videos of attack attempts of 50 individuals, more details about the dataset is presented in Section 4.2. The speed and complexity of the developed algorithms was not considered. Every participant was supposed to deliver a score after processing 230 frames of each video in the test data set. In the following a short overview of the algorithms of participants is presented.

The Ambient Intelligence Laboratory (AMILAB) team combined all three indicators of anti-spoofing attacks (motion, texture and liveness). They analyzed each frame of video and computed a confidence score of being real or fake. In the process, they extracted different types of visual features (color, texture, edges) and used a set of SVMs for the calculation of confidence score. Two more sources of classification scores were used, one is based on the average movement for which motion detection technique [LHGT03] is used. The other source uses eye blink detection [CB05]. The scores are combined by the means of the Dynamic Score Combination Methodology [TGR09] that dynamically find the weights of scores. After this step they analyze liveness degree of the person. This team achieved good results with only one false rejection on the test set in the competition.

The research group of the Institute of Computing UNICAMP, Brazil, combined in their approach all three indicators. They extracted robust set of low level features from regions within the faces. These are descriptors based on the histogram of oriented gradients (HOG) [DT05], Gray Level Co-Occurrence Matrix (GLCM) [cit73] and histograms of shearlet coefficients (HSC)[SdSDP11]. The feature descriptors are weighted with partial least square regression (PLS). To exploit those features, video sequences are divided into parts, faces are detected, cropped and rescaled.

IDIAP team used only texture indicator in their approach. At first, videos are converted into gray-scale, then local binary patterns (LBP) operator (with R=1 and 8 neighbor pixels) is applied to every single image in the stream, and as a result, a global histogram is build with 256 bins where the outputs from LBP operator (2D) are accumulated [TPM02]. This global histogram is compared with a reference histogram generated and trained by using all available fake videos in the training set.

The approach proposed by CASIA team was based on three observations that build the basis for the classifiers used in the challenge.

- Classifier 1 - Non-rigid Motion. The idea here is that the real videos have non-rigid

[1] http://www.idiap.ch/dataset/printattack

motions, especially in the mouth and eyes, while printed photos have only rigid motions (translation, scaling and rotation). Therefore, they can be well modeled by this transformations. For the detection of non-rigid motion in video, the research group used the RASL technique proposed in [PGW+] that is able to align a series of affine transformed images (spoofing videos). Hence, the classifier is trained on the differences of aligned frames.

- Classifier 2 - Noise. This classifier is built on the noise differences between real and spoof videos. According to the authors, the spoofed videos have usually more noise. Noise in images is evaluated by Haar wavelets.
- Classifier 3 - Face-Backgroound Dependency. This classifier is based on the idea that the motion of head in real video is independent of the background motion. In the spoof video the small area of background around the face is moving with the face (printed image). Therefore, this classifier is trained on the ratio of motion in the face and background.

The fusion of the three classifiers depends on the background. If the background is complex, the classifier 1 and 3 are fused for score computation. If the background is uniform, classifier 1 and 2 are fused. This approach achieved good results and appears to be robust, but this comes with high complexity cost.

The approach presented by UOULU team relies on the detection of printing quality defects using texture information. This approach only considers single frames. This is achieved by applying Sobel filter followed by the calculation of LBP features histogram and the classification decision is made by SVM classifier.

The SIANI team used the evolution of face appearance and location as a liveness clue. The motion of certain areas (face, non-face, mouth and eyes) between frames is analyzed. Classification decision is built by a Bayesian Network based on the motion statistics. The results of the six approaches participating in the competition [CAM+11] are shown in Table 2.

4 Discussion

In this section, the expected performance and challenges facing solutions based on different types of liveness indicators are discussed. This is linked to different possible spoof attack scenarios. In this framework, a comparison between the different discussed works is presented. Moreover, the main used datasets are discussed.

4.1 Liveness Indicators and Spoof Attacks

In this work, liveness detection approaches are categorized based on the type of liveness indicator (clue) used to assist the liveness of 2D faces. Three main types of indicators

where introduced, motion, texture and life sign. Each of those indicators represents certain clues of the liveness of the face. Each can act as anti-spoofing measure in certain scenarios. However, each can be prone to certain spoofing situations. Here, the three types of liveness indicators are discussed and their advantages and disadvantages in different attack scenarios are presented.

Motion analysis tries to differentiate the motion pattern between 3D and 2D faces. The assumption here is that real (live) faces are 3D structures, while spoof attack faces are 2D images. Those images can be printed on paper or displayed on screens. Motion analysis usually depends on optical flow calculated from video sequences.

Texture analysis identify real (live) faces by having different texture qualities than printed face images. This approach assumes that fake faces are printed on paper, and the printing process and/or the paper structure produce texture features that can differentiate those printed images from real face images.

Detection of life signs can be categorized itself into two categories. First that assumes certain known interaction from the user (subject to be verified). In this situation the user performs a certain task to verify the liveness of his face image. This task can be a certain move that can be considered as a challenge response or a motion password. Users who will perform their task correctly are assumed to be real. The second category does not assume interaction from the user, but focuses on certain movements of certain parts of the face, such as eye blinking, and will consider those movements as a sign of life and therefore a real face.

The selection of liveness indicator type to build a liveness detection solution around is directly related to the expected application scenario and spoof attacks. In the following, different attack scenarios are assumed and the expected performance of the different liveness indicators is discussed.

The usually assumed face spoof attacks are the printed face attack. Here, the user face is printed on a paper and presented in front of the camera for verification or identification. Using texture analysis to identify real faces is possible in this situation, as the printing procedure and paper usually contains high texture characteristics. Nonetheless, very diverse paper and printing textures can occur, and the systems built on texture analysis must be robust to different texture patterns. This requires the existence of a very diverse dataset. It is also possible that the attack is performed using a photo displayed on a screen, in this case very low texture information is in hand. Using motion analysis or life sign detection approaches will not be challenged by the diversity of texture information as it does not depend on the texture but rather on movement information.

As mentioned above, using motion analysis as liveness indicator will help get over the dependency on certain texture patterns. However, motion analysis may face problems when low motion information is in hand. This can happen because of the behavior of the user, high noisy images and low resolution. Motion analysis might also failed when spoof attacks is performed using more sophisticated methods, just like 3D sculpture face model.

One the other hand, life sign detection based on the movement of certain face parts can produce a robust system that is independent on the texture information and is secure against attacks using 3D face models. However, this method may face challenges detecting the

Work	Motion	Texture	Life sign	Dataset	Result
AMILAB*	X	X	X	PRINT-ATTACK	FAR=0.00 FRR=1.25
CASIA*	X	X	X	PRINT-ATTACK	FAR=0.00 FRR=0.00
IDIAP*		X	X	PRINT-ATTACK	FAR=0.00 FRR=0.00
SIANI*	X			PRINT-ATTACK	FAR=0.00 FRR=21.25
UNICAMP*	X	X	X	PRINT-ATTACK	FAR=1.25 FRR=0.00
UOULU*		X		PRINT-ATTACK	FAR=0.00 FRR=0.00
[kJuJhY06]			X	100 sampels, 10 objects	FAR=0.01 FRR=0.08
[BLLJ09]	X			100 samples, 10 objects	betw. 83% - 100%
[KFB]	X			XM2VTS database, 200 live and 200 fake	error rate 1% by 400 test sequences
[KFFB]			X	XM2VTS database + simulated scenarios	succesfull recog 0.73 for 100 persons

Table 2: Overview of the discussed approaches. *[CAM+11].

certain motions considered.

Life sign detection based on user interaction and performing specific tasks can produce very robust and secure anti-spoof solutions. This solution does not depend on the texture information and is respectively robust to image properties. However, this solution requires the collaboration of the user and is not suitable for ubiquitous systems.

Table 2 presents a comparison between the different discussed works. This table presents the different approaches and the liveness indicator used. It also shows the dataset used for development and evaluation and the performance reached. One can notice the high accuracy achieved when using dataset (scenario) that matches the liveness indication used. This is clear with the use of, the texture rich, PRINT-ATTACK dataset and texture based solutions. Moreover, one can notice the high performance of solutions combining more than one liveness indicator type. Those solutions are also expected to have higher robustness to different spoof attacks. Table 3 lists a brief comparison between the different liveness indicators.

4.2 Datasets

In order to develop and test a face liveness detection solution, an informative and diverse dataset that imitate the expected application scenarios is needed. The application scenario and expected spoof attacks are linked with the developed solutions through the type of information extracted from the images and used for liveness detection (liveness indicators).

Two publicly available datasets were mainly used for the development of the presented face liveness detection solutions. One is the more recently available PRINT-ATTACK [AM11] dataset that was explicitly built to develop and evaluate face liveness detection

Liveness Indicator	pros	cons
Texture	simple implementation, good results in known scenarios, possible decision from one frame, no user-collaboration needed	needs data that covers all possible attacks, problem with low textural attacks
Motion	very hard to spoof by 2D face image, independent of texture, no user-collaboration needed	needs video sequence, can be spoofed by 3D sculptures, needs high quality image, challenged by videos with low motion activity
Life sign	very hard to spoof by 2D face images or 3D sculpture, independent of texture	may need collaboration from user, depends on landmark detection in the face, needs video sequences

Table 3: Overview of the discussed approaches

approaches that assumes an attack by printed face images. The other is the XM2VTS dataset [MMK+99], this dataset was initially built for the development and evaluation of person identification and verification based of frontal and profile faces, as well as, speech analysis.

The PRINT-ATTACK database consists of 200 videos of printed face photo attacks and 200 videos of real access attempts. The videos where collected from 50 different persons under different lightning conditions. Videos are captured by having a real person or a printed photo trying to access a laptop through a webcam. The color videos sequences where at least 9 seconds each and were taken under two illumination scenarios, a controlled scenario and a diverse scenario with more complex background and natural light.

Attack videos where recorded under two conditions. First, the videos where taken of printed face photos hold still on a stand. A second attack set was captured while the attacker holds the face image by hand. The data was split into four groups, training, development, testing and enrolment.

One can notice the highly visible texture patterns in the attacks paper. This might produce high liveness detection accuracy when the solutions based on texture features are tested, as one can see in table 2. The non-interactive nature of this dataset makes it not suitable for developing interactive liveness detection solution but suitable for detection of life signs under realistic conditions as it contains no intended movement. The structure of the dataset is very helpful for developing motion analysis based liveness detection solutions. As the data set contains two types of attacks, fixed and hand held pictures, this help analyzing the differences of motion patterns between real faces and the different attacks.

The XM2VTS dataset was used in some works to develop face liveness detection solutions [KFB] [KFFB]. This dataset was initially developed for person identification based on speech analysis, frontal faces and profile faces. The dataset contains data collected from 295 individuals.

Two parts of the dataset are interesting for the face liveness detection solutions. First, the

rotation shots. Those video sequences capture the rotation movement of individuals heads. The second part is the speech shots. Those video sequences capture individuals reading three predefined sentences. Spoof attacks based on this dataset can be built by printing a certain frame of the sequences and capture a video of it in a spoof attack scenario [KFB].

This dataset is suitable to be used in developing liveness detection that is based on life signs detection especially when interactive users are expected [KFFB]. On the other hand, as this dataset was not explicitly built for liveness detection, it does not contain a standard subset of spoof attack video sequences.

Many Self-captured datasets were used for the development and evaluation of liveness detection algorithms [kJuJhY06][BLLJ09].

For the future development of public databases, the three types of liveness indicators (information) should be considered. Non-interactive video sequences for motion analysis and non-interactive life sign detections must include along with a set of interactive sequences where the users performs certain tasks. Future attack datasets must consider the weaknesses of the available solution. Those weaknesses are limited representations of possible textures, static or unrealistic movement of printed images attacks and their representation, and the possibility of more sophisticated attacks such like 3D sculpture faces.

5 Conclusion

This work provided an overview on the problem of 2D face liveness detection. It presented a categorization scheme for the possible solutions based on the type of liveness indicator used. A review of diverse solutions was presented and discussed in the framework of the presented categories. A discussion was built to link the possible attack scenarios, the current and future solutions, as well as the available and future datasets. This work aimed to provide basic framework for understanding the problem of face liveness detection, the state of the art solutions, and the future challenges under different spoof attack scenarios.

References

[AM11] A. Anjos and S. Marcel. Counter-Measures to Photo Attacks in Face Recognition: a public database and a baseline. 2011.

[BFB94] John L. Barron, David J. Fleet, and Steven S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[BLLJ09] Wei Bao, Hong Li, Nan Li, and Wei Jiang. A liveness detection method for face recognition based on optical flow field. In *Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on*, pages 233 –236, april 2009.

[CAM+11] M. M. Chakka, A. Anjos, S. Marcel, R. Tronci, D. Muntoni, G. Fadda, M. Pili, N. Sirena, G. Murgia, M. Ristori, F. Roli, J. Yan, D. Yi, Z. Lei, Z. Zhang, S. Z.Li, W. R. Schwartz, A. Rocha, H. Pedrini, J. Lorenzo-Navarro, M. Castrillón-Santana,

J. Maatta, A. Hadid, and M. Pietikainen. Competition on Counter Measures to 2-D Facial Spoofing Attacks. In *Proceedings of IAPR IEEE International Joint Conference on Biometrics (IJCB), Washington DC, USA*, October 2011.

[CB05] Michael Chau and Margrit Betke. Real time eye tracking and blink detection with USB cameras. Technical report, 2005.

[cit73] Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(3):610–621, November 1973.

[DT05] Navneet Dalal and Bill Triggs. Histograms of Oriented Gradients for Human Detection. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01*, CVPR '05, pages 886–893, Washington, DC, USA, 2005. IEEE Computer Society.

[KFB] K. Kollreider, H. Fronthaler, and J. Bigun. Evaluating Liveness by Face Images and the Structure Tensor.

[KFFB] Klaus Kollreider, Hartwig Fronthaler, Maycel Isaac Faraj, and Josef Bigun. Real-Time Face Detection and Motion Analysis With Application in Liveness Assessment.

[kJuJhY06] Hyung keun Jee, Sung uk Jung, and Jang hee Yoo. Liveness Detection for Embedded Face Recognition System. *International Journal of Biomedical Sciences*, 2006.

[LHGT03] Liyuan Li, Weimin Huang, Irene Y. H. Gu, and Qi Tian. Foreground object detection from videos containing complex background. In *Proceedings of the eleventh ACM international conference on Multimedia*, MULTIMEDIA '03, 2003.

[MMK+99] K. Messer, J. Matas, J. Kittler, J. Lttin, and G. Maitre. XM2VTSDB: The Extended M2VTS Database. In *In Second International Conference on Audio and Video-based Biometric Person Authentication*, pages 72–77, 1999.

[PGW+] Yigang Peng, Arvind Ganesh, John Wright, Wenli Xu, and Yi Ma. RASL: Robust Alignment by Sparse and Low-rank Decomposition for Linearly Correlated Images â.

[SdSDP11] William Robson Schwartz, Ricardo Dutra da Silva, Larry S. Davis, and Hlio Pedrini. A novel feature descriptor based on the shearlet transform. In *ICIP*. IEEE, 2011.

[TGR09] Roberto Tronci, Giorgio Giacinto, and Fabio Roli. Dynamic Score Combination: A Supervised and Unsupervised Score Combination Method. In *Proceedings of the 6th International Conference on Machine Learning and Data Mining in Pattern Recognition*, MLDM '09, pages 163–177, Berlin, Heidelberg, 2009. Springer-Verlag.

[TPM02] Ojala T. and Menp T. Pietikinen M. Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns. 2002. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7):971 - 987.

[VJ01] Paul A. Viola and Michael J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. In *CVPR (1)*, pages 511–518, 2001.

[WLW04] H. Wang, S. Z. Li, and Y. Wang. Face recognition under varying lighting conditions using self quotient image. In *Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition*. IEEE Computer Society, 2004.

[ZSBF08] X. Zhou, H. Seibert, C. Busch, and W. Funk. A 3D Face Recognition Algorithm Using Histogram-based Features. In Stavros J. Perantonis, Nickolas S. Sapidis, Michela Spagnuolo, and Daniel Thalmann, editors, *3DOR*. Eurographics Association, 2008.