Evaluating liveness by face images and the structure tensor

Klaus Kollreider, Hartwig Fronthaler and Josef Bigun

Kollreider K, Fronthaler H, Bigun J. Evaluating liveness by face images and the structure tensor. In: Fourth IEEE Workshop on Automatic Identification Advanced Technologies, 2005. IEEE; 2005. p. 75-80.

DOI: http://dx.doi.org/10.1109/AUTOID.2005.20

Evaluating Liveness by Face Images and the Structure Tensor

K. Kollreider, H. Fronthaler and J. Bigun
Halmstad University, SE-30118, Sweden

***Abstract*— *A technique evaluating liveness in short face image sequences is presented. The intended purpose of the proposed system is to assist in a biometric authentication framework, by adding liveness awareness in a non-intrusive manner. Analysing the trajectories of single parts of a live face reveal valuable information to discriminate it against a spoofed one. The proposed system uses a lightweight novel optical flow, which is especially applicable in face motion estimation based on the structure tensor and a few frames. It uses a model-based local Gabor decomposition and SVM experts for face part detection. An alternative approach for face part detection using optical flow pattern matching is introduced as well. Experimental results on the proposed system are presented.*

I. Introduction

Liveness Detection is a highly desirable, yet rather unexplored anti-spoofing measure in biometric identity authentication [20], [21]. Especially in face analysis only a few approaches address this subject: In [7] a depth map is constructed by recovering 3D structure from motion. This depth map is constant in case of a photograph, even if in motion, whereas a live face yields varying depth values. Unlike in our work motion is estimated by a feature based method. Another way is to analyse the frequency spectrum of a live face [17], defining two descriptors to measure the high frequency proportion and the temporal variance of all frequencies. Their method relies on both the lack of quality of a photograph and the change of mimics and pose in a live face.

In our approach we combine face part detection and optical flow estimation to determine a liveness score. The unique trajectory of face parts in case of a live face sequence is exploited to discriminate it against a spoofed one. We use an adapted optical flow termed Optical Flow of Lines (OFL). It is inspired by optical flow approaches, which are able to differ between motion of points and motion of lines (e.g. Tensor approach [6]). As the name suggests, it is specialized on motion of lines only. Requiring only 3 images the OFL approach is a lightweight, energy based OF method, entirely realized employing 2D Gabor filters [3], [11], [15]. A detailed review of existing optical flow techniques is given in [1].

For face part detection we combine OF pattern matching with a model-based technique [4] employing Gabor features on a log-polar grid [22], [23] and SVM. Gabor filters are a class of powerful face recognition features [9], [16], [23] which have impulse responses resembling those of simple cells in visual cortex [13], [19].

II. Basic Strategy

Obviously even the best authentication methods will be weak if they cannot distinguish between a photograph and the live presence of a client. An underestimated problem in face authentication studies is the claim of someone else's identity by using a high quality photograph, whether in motion or not. Essentially three possibilities to make such a system liveness aware can be identified:

1) Deploying an excessive and costly system configuration (e.g several cameras including stereo, heat sensitive cameras, etc.).
2) Interacting with the client (e.g. Automated Teller Machine) demanding realtime responses (e.g. talk, blink, etc.) and/or collecting multiple biometrics [2], [4], [14]. These measures demand active collaboration, which is complementary to our study. They may not be employable for a variety of reasons including the nature of the application and a perception of being intrusive.
3) Exploiting the motion characteristics of a 3D face by using an image sequence. In order to bypass client interaction, preferably few images containing some natural motion (behaviour) should be sufficient for reliable liveness detection.

In a biometric authentication framework the proposed system is a single component pursuing liveness detection in accordance with the last cathegory above. It analyses a face image sequence captured by one camera and delivers a probability, whether it detected a face and whether it is live.

Fig. 1. Example face image sequence

The basic idea relies on the assumption that a 3D face generates a special 2D motion which is higher at central face parts (e.g. nose) compared to the outer face regions (e.g. ears). Ideally, in terms of liveness detection, the outer and the inner parts move additionally in opposite directions. This case is visualized in figure 1 and 2 respectively, where a head slightly rotates to the left (from the person's view). Figure 2 shows the horizontal OFL (optical flow estimate in horizontal direction only) from the image sequence displayed in figure 1. The

rectangles indicate the focused face parts and their motion (note the signs).
In other words, parts nearer to the camera move differently to parts which are further away in a live face. This assumption requires at least partially rotational movement of the head, which we consider as natural and unintentional human behavior. On the contrary, a translated photograph generates constant motion at various face regions.

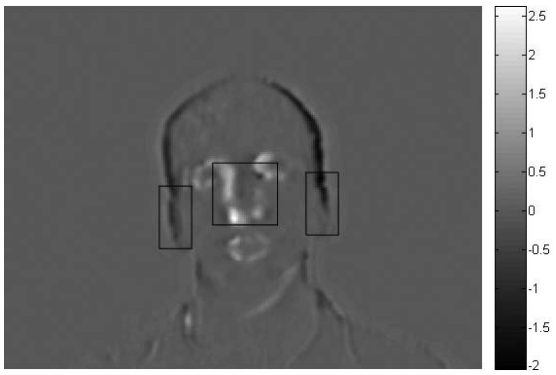

Fig. 2. Horizontal OFL with rectangles indicating focused regions

In order to exploit these characteristics, we utilize an optical flow estimation and a face part detection. For the latter we employ a model-based Gabor decomposition [4], but we also present an intuitive approach by OF pattern matching. Knowing the face parts' position and comparing how fast they are moving relative to each other and into which directions, enables us to discriminate a live face against a photograph.

III. Optical Flow Estimation

The subsequently presented Optical Flow of Lines (OFL) relies on some assumptions and simplifications. First, as hinted, the OFL approach can only handle motion of lines, often referred to as normal motion. Second, it assumes lines to be either horizontal or vertical when estimating the velocity components. We motivate these simplifications by regarding lines, and especially those horizontally and vertically oriented, as the dominant structure in a face image of known scale range. We assume these features to be sufficiently robust for spatiotemporal analysis. This allows, as shall be seen, to reduce a 3D-minimization problem to 2D.

A. Theoretical Approach

In the general case of parallel lines undergoing a constant motion, parallel planes are generated in the spatiotemporal room, having a common normal unit vector $\hat{k} = (k_x, k_y, k_t)$ to describe them. The tilt α of this normal vector with regard to the xy-plane corresponds to the absolute velocity of the concerned lines in 2D, which is also stated in equation 1.

$$|v| = \tan\alpha = \frac{|k_t|}{\sqrt{k_x^2 + k_y^2}} \tag{1}$$

Due to the aperture problem, we can only determine the normal optical flow, which is in the spatial direction of $\hat{k}$. The horizontal and vertical velocity components are denoted in equation 2.

$$v_x = -\frac{k_x \cdot k_t}{k_x^2 + k_y^2} \qquad v_y = -\frac{k_y \cdot k_t}{k_x^2 + k_y^2} \tag{2}$$

As a consequence, we need to estimate $\hat{k}$ in order to calculate the normal velocity, which corresponds to orientation estimation in 3D. This is further equivalent to an eigenvalue analysis of the multidimensional structure tensor. This allows the minimization process of fitting a line or a plane to be carried out without actually Fourier transforming [6]. Applied to optical flow estimation, known as tensor method, the eigenvector belonging to the solely large eigenvalue of the 3D structure tensor directs into the direction of $\hat{k}$, if normal motion is detected (spectral energy is concentrated to a line). If we assume to have extracted vertical and horizontal lines separately throughout the image sequence (directional filtering), and we are interested in their translation only, we can formulate the optical flow in a spatially separable manner. We are then looking for tilted planes, which stay parallel to either one spatial axis. The determination of these tilts α corresponds to 2D orientation estimation along the parallel axis, which is in xt- and yt-images. The 2D normal vectors $\hat{k}_1$ and $\hat{k}_2$ of the planes parallel to the y- and x-axis respectively are related to $\hat{k}$ considering respectively k_y and k_x zero. This is also valid for the the horizontal and vertical velocity components, which reduce to:

$$v_x = -\tan\alpha = -\frac{k_t}{|k_x|} \qquad v_y = -\tan\alpha = -\frac{k_t}{|k_y|} \tag{3}$$

The optical flow estimation is thus determined by orientation in two dimensions for each component.
In 2D orientation estimation, eigenvalue analysis of the corresponding 2D structure tensor can be replaced by averaging both the square of a complex valued gradient image and its absolute value [6]. If we denote Fx and Fy as the image sequences containing extracted vertical and horizontal lines respectively, we can establish the relationships as in equation 4 and 5.

$$\hat{k}_1{}^2 = (k_x + i \cdot k_t)^2 = \int\int \left(\frac{\partial Fx}{\partial x} + i \cdot \frac{\partial Fx}{\partial t} \right)^2 dx\, dt = V_x \tag{4}$$

$$\hat{k}_2{}^2 = (k_y + i \cdot k_t)^2 = \int\int \left(\frac{\partial Fy}{\partial y} + i \cdot \frac{\partial Fy}{\partial t} \right)^2 dy\, dt = V_y \tag{5}$$

This corresponds to linear symmetry detection in xt- and yt-images respectively. The complex numbers V_x and V_y directly encode the optimal direction in double angle representation and the error [12]. In other words, $\arg V_x$ and $\arg V_y$ equal 2α in equation 3, leading to the estimated velocity components stated in 6.

$$v_x = -\tan\left(\frac{1}{2}\arg V_x\right) \qquad v_y = -\tan\left(\frac{1}{2}\arg V_y\right) \tag{6}$$

As we have assumed the image sequences to contain lines only, we neglect a possible orientation estimation error due to noise.

B. Implementation

The implementation of the OFL follows the concepts above, using Gabor filters for all tasks. A filter bank comprising 3 Gabor filters each tuned to one orientation at a common absolute frequency is designed. In the following we also refer to these filters as G_l with $l = 1, 2, 3$, labelling their orientations $0°$, $60°$ and $120°$. The center frequency of each filter is subsequently referred to as z_l, encoding the angular frequency components as complex numbers. The filters are applied on xt- and yt-images and need to be small in our case (using only 3 images). The input images are referred to as Im_l. In the following, a summary to calculate the component velocity v_x is presented. Its adaptation to calculate v_y is achieved analogously.

1) Filter each Im_l with G_1 to get Fx_l
2) Form a 3D space-time image stack out of Fx_l, and slice it along the vertical axis to receive all xt-images, referred to as $xtIm_{1-c}$ (c being the size of any input image in vertical direction)
3) Filter each $|xtIm_{1-c}|$ with G_1, G_2, G_3 separately, giving $fxtIm_{1-c,1-3}$
4) Calculate $V_{\mathbf{x}}$ at every pixel of the center input image (Im_2) by:

$$V_x(x,y) = \sum_{i=1}^{3} |fxtIm_{y,i}(x,2)|^2 \cdot {z_i}^2 \qquad (7)$$

5) Calculate v_x out of V_x:
$v_x(x,y) = -\tan(0.5 \cdot \arg V_x(x,y))$
6) Consider $v_x(x_i, y_i)$ only if
$|v_x(x_i, y_i)| < \tau_1$ and $|\Im(Fx_2(x_i, y_i))| > \tau_2$

When calculating v_y we rotated Im_l by $90°$ before step 1. Consequently the component velocity matrix has been rotated by $-90°$ after step 6. Equation 7 is similar to the scheme of quadrature filters suggested by [15] although our filters are different. The idea is to weight a direction z_l normal to the filter's selective orientation by the filter response. The vector sum of the three weighted arrows in double angle representation yields an estimate of the observed linear symmetry (see equation 7). Alternatively the direct approach described in [5] could be used for orientation estimation. In step 6, two constraints are evaluated. First, the maximum reliable tilt angle α is bounded, in other words, the maximum component velocity is restricted by τ_1. Second, to ensure that the component velocity is only considered at reliable vertical/horizontal line structure, we exploit the edge filter property of G_1 from the directional filtering done in step 1: We demand the imaginary part of the filter response to be greater in magnitude than τ_2.
The horizontal and vertical component velocity at each point is combined to a complex image OF_{im}, having v_x in its real part and v_y in the imaginary part, as described in equation 8.

$$\underset{x,y \in Im_2}{OF_{\text{im}}} = v_x(x,y) + i \cdot v_y(x,y) \qquad (8)$$

The OFL is applied on two test sequences. The results are shown in figure 3.

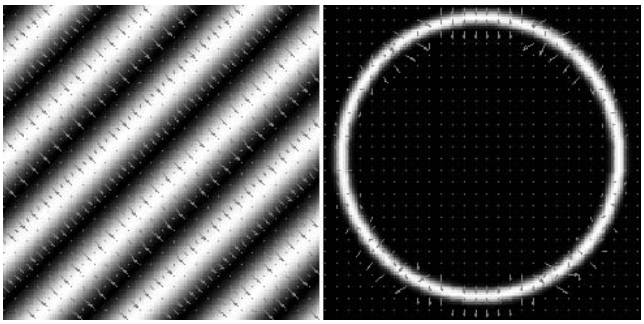

Fig. 3. OFL applied on sinusoid - and circle image sequence

The left image displays the center frame of a sinusoid image sequence, undergoing a diagonal top to bottom translation by 2 pixels/picture. The right image shows the center frame of a circle image sequence, exposed to a vertical top to bottom movement by 2 pixels/picture. In both images, the flow arrows estimated by our approach are superimposed. As can be observed, by considering the horizontal and vertical portions of off-axis lines, the OFL is accurate at perfectly horizontal and vertical lines and reasonably accurate at oblique directions.

IV. Face Part Detection

We observe the trajectory of three facial regions: Searching for the face center, we concentrate on the nose and the eyes whereas in case of the two side models the facial curvature and the ears are focused on. To reliably detect these face parts we combine optical flow pattern matching and model-based Gabor feature classification.

A. Optical Flow Pattern Matching

The face center can be approximated by reusing information from the optical flow estimation, because the region around the eyes and the nose shows a characteristic flow pattern. A template containing the flow information of an average face center is created offline, visualized in figure 4.

Fig. 4. Example face center template for OF pattern matching

The first image in figure 4 displays $|v_x|$ around the face center of a horizontal only movement whereas the second one shows $|v_y|$ taken from a vertical only motion. The complete template OF_{temp} used for the first matching is complex, combining v_x and v_y analog to equation 8. Its absolute value is displayed in the third image of figure 4. The template has approximately 2-3% the size of OF_{im}. We start by calculating

the absolute similarity of OF_{temp} at each pixel of OF_{im} using the Schwartz inequality (normalized correlation),

$$\frac{|< OF_{\mathrm{im}}, OF_{\mathrm{temp}} >|}{\|OF_{\mathrm{im}}\| \cdot \|OF_{\mathrm{temp}}\|} \leq 1 \tag{9}$$

resulting in a similarity matrix sim with values in $[0, 1]$. The $max(sim)$ is stored in f_{cer} and two further similarities are calculated at that position:

$$-1 \leq \frac{< \Re(OF_{im}), \Re(OF_{temp}) >}{\|\Re(OF_{im})\| \cdot \|\Re(OF_{temp})\|} \leq 1 \tag{10a}$$

$$-1 \leq \frac{< \Im(OF_{im}), \Im(OF_{temp}) >}{\|\Im(OF_{im})\| \cdot \|\Im(OF_{temp})\|} \leq 1 \tag{10b}$$

Equations 10a and 10b give a scalar in $[-1, 1]$ referred to as sim_h and sim_v respectively. These similarity measures indicate the directions (by their signs) in which the face center moves and the relative velocities, e.g. whether the actual movement is more a horizontal than a vertical one.
The procedure described in this section can effectively be employed to detect the center facial region, whereas it is less applicable to detect the boundary facial features. Furthermore it is a quickly computed origin for subsequent analysis.

B. Model-based Gabor Feature Extraction

The second method presented for face part detection is similar to [4]. Essentially, features are extracted at certain points of a non-uniform retinotopic grid in order to measure image properties of a face [23], [24]. These features constitute models, which are classified by trained SVM experts [8].

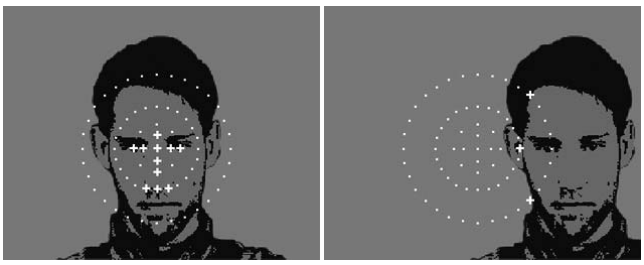

Fig. 5. Models for center and right hand side facial region

Each model uses specific points chosen out of the retinotopic grid, which are marked by plus signs in figure 5. The center points are denser and aligned to the structure in the middle of a face, whereas the outer model points are chosen to be bent along the facial curvature. At these points, features are extracted by means of a series of Gabor filters, which in our case consist of five frequency - and six orientation channels. The filter bank is designed in the log-polar domain (a logarithmically scaled polar space), where the Gabor filters are uniformly distributed Gaussian bells [3]. This ensures that the designed Gabor filters evenly cover the Fourier domain. Only specific frequency - and orientation channels are used within the models, because the approximate distance range of the faces was assumed to be known. Like in the OFL approach, where we concentrate on horizontal and vertical lines, we select the Gabor filter orientations accordingly. The selected frequency channels are tuned to the scale of the features we are interested in. This adaption is done to focus on few but important features. The classification performance generally deteriorates when the number of features increases, if no dimension reduction is done e.g. due to restrictions such as realtime operation, [4], [10].
The feature vector $\vec{k}$ for one point of a specific model (e.g. a plus sign in figure 5) consists of the absolute value of single scalar products between the image and the chosen Gabor filters at that point. The dimension of $\vec{k}$ equals the product of the employed amount of frequencies and orientations.
The complete feature vector $\vec{x}$, contains the elements of $\vec{k}$ for each of the concerned grid points, ordered by the latter. It is used to train the corresponding SVM classifiers (RBF kernel), and consequently serves as input to the trained experts. To add face generalisation, the three classifiers (center, right, left) are trained with the last 145 frontal face images from the XM2VTS Database [18]. All images are downsized to a resolution of 300x240 to reduce the computing time. The three face part experts deliver the certainty measures g_{cer1-3}, where subscript 1 corresponds to the center (inner face part) and 2,3 to the right and left hand-side (outer face parts) respectively. If the optical flow pattern matching method described previously is preceding the Gabor feature based method, the critical search area for the latter can significantly be reduced.

V. Liveness Detection

The flow chart in figure 6 shows all components of the liveness detection algorithm and their interconnection.

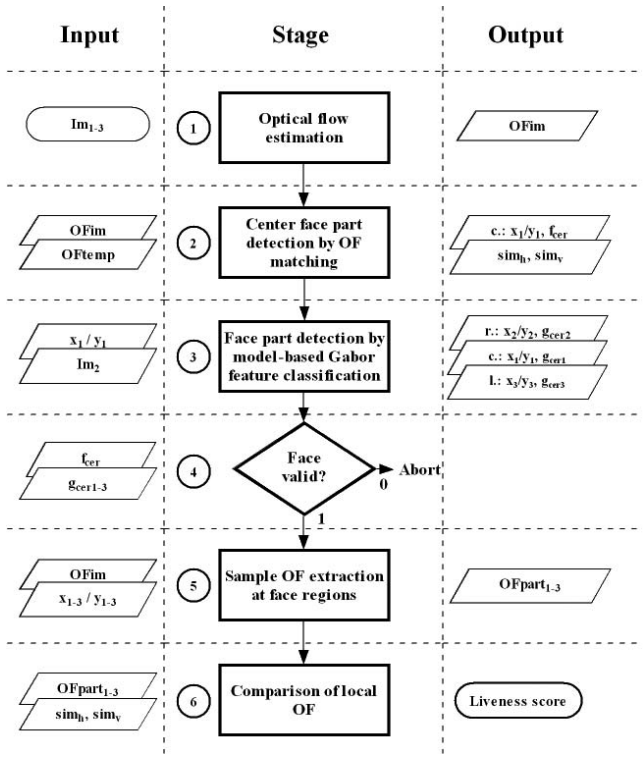

Fig. 6. Flow chart of Liveness detection

The following enumeration refers to the stage numbers in figure 6 and details each step:

1) The OFL at Im_2 is calculated using the algorithm described in subsection III-B.
2) The face center is detected using OF pattern matching as described in chapter IV-A.
3) All three face parts are detected in the center image Im_2 by model-based Gabor feature classification, outlined in subsection IV-B. As the position of the face center is already known from stage 2, the feature extraction is only employed in a small neighborhood. The two outer models are extracted each in the expected neighborhood to the right and left of the face center. The purpose is to confirm the face center position detected rapidly in stage 2 and to ensure the presence of an actual face.
4) In this step, the certainties f_{cer} and g_{cer1-3} are reviewed. If one of them is below a common threshold τ, the sequence is regarded to be unsuitable due to an insufficiently recognizable face (non-face), yielding a liveness score of 0.
5) A rectangular area around each face part's (central) position x_{1-3}/y_{1-3} is cut out of OF_{im} and stored as image parts $OF_{part1-3}$. Figure 2 indicates examples for these regions.
6) Finally $OF_{part1-3}$ are compared relatively to each other. Only the values of each OF_{part} greater than half its maximum absolute value are considered. The remaining values are divided by their total number preparing mean value calculation. We decide to concentrate on the primary movement only. The ratios cr and cl, which contribute to the final liveness score are calculated as follows:

   if $|sim_h| > |sim_v|$

   $$cr = \frac{\sum \Re(OF_{part1})}{\sum \Re(OF_{part2})}, \qquad cl = \frac{\sum \Re(OF_{part1})}{\sum \Re(OF_{part3})}$$

   else

   $$cr = \frac{\sum \Im(OF_{part1})}{\sum \Im(OF_{part2})}, \qquad cl = \frac{\sum \Im(OF_{part1})}{\sum \Im(OF_{part3})}$$

   Depending on the primary movement, the ratios cr and cl compare the real or the imaginary part of OF_{part1} with OF_{part2} and OF_{part3} respectively. A ratio between the center and a side motion having an absolute value greater than 1, indicates liveness. Furthermore, negative ratios suggest oppositely moving face parts. The final liveness score is then constructed as,

   $L = \frac{1}{4}[(|cr| > 1) + (|cl| > 1) + (\ominus cr < 0) + (\ominus cl < 0)]$

   where $\ominus x$ means "sign of" x. The score is an value in [0,1], where 0 indicates no liveness and 1 represents the maximum liveness.

VI. Experimental Results

We use the "Head Rotation Shot"-subset (DVD002 media) of the XM2VTS database to evaluate the performance of our liveness detection scheme. We employ the first 100 videos from the first session only. All data is downsized from 720x576 to 300x240 pixels.

A. Live and Non-live Sequences

For the "live sequences", which are expected to gain a high liveness score, 2 sequences (of 3 frames each) containing partial rotation are cut from each video. The possible movements (to the left, to the right, up, down) are evenly present in the sequences.

On the other hand, "non-live sequences" have to be manufactured because a database for playback attacks using photographs is not available: For each person one frame is taken out of a respective live sequence, and translated twice each horizontally and vertically to produce sequences of 3 frames. This yields 2 non-live sequences per person, imitating high resolution photographs in motion.

Two live sequences on top of their non-live counterparts are displayed in figure 7. The first two rows contain a horizontal movement, whereas it is a vertical one in the last two rows. Note that the motion in the bottom live sequence is hardly noticeable by the human eye.

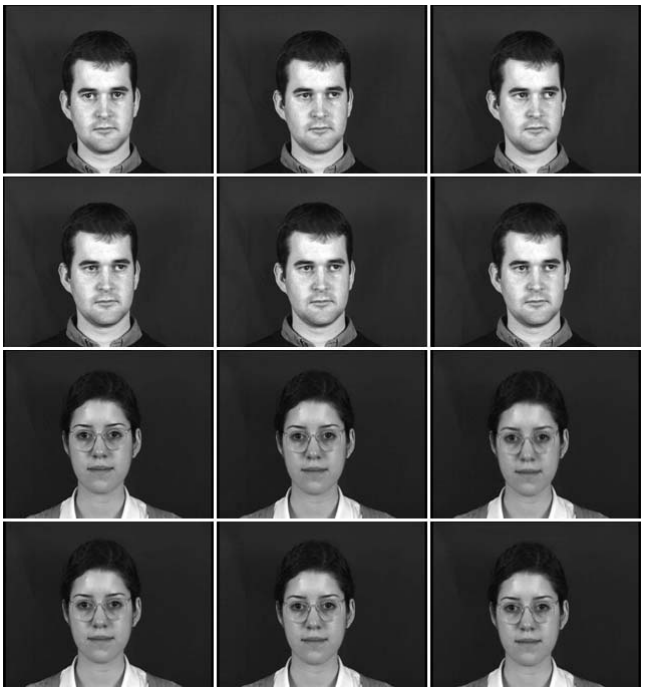

Fig. 7. Row 1, 3: live sequences; Row 2, 4: non-live sequences (playback attack)

B. Test Results

A total of 200 live - and 200 non-live sequences were analysed by the liveness detection system described in the previous chapter. The liveness score achieved by each sequence was stored as primary result. In addition to that, the progression of the face part detection was monitored for each sequence. The Liveness score distribution for both live and non-live is detailed in table I.

As can be observed, most of the non-live sequences scored 0, whereas most of the live sequences achieved a score of 0.75. No live sequence scored below 0.5. The explanation for non-live sequences achieving a score greater than 0 lies in up/down-movement, when there is less horizontal line structure to observe at the side parts of a face, compared to the face center. If there are no lines present, or only few

Liveness score	# Non-live seq.	# Live seq.
0	148	0
0.25	49	0
0.5	3	38
0.75	0	120
1	0	42

TABLE I

LIVENESS SCORE DISTRIBUTION FOR LIVE - AND NON-LIVE SEQUENCES

sufficiently inside a checked side region, the velocity at the face center is admittedly measured to be higher. This happened only at one side region at a time, except for three sequences of the test data. It has also been observed, that eyeglasses can lower the liveness score of live sequences, as they are near the camera even at the sides. However, as the table suggests, by considering a sequence achieving a liveness score ≥ 0.5 as live, the system separates the 400 test sequences with an error rate of 0.75%. The system is error-free if only sequences containing horizontal movement are considered

The success rate of the alternative face center detection (by OF patterns) was further encouraging, as it was 92% in isolation. The main reasons for false localization were eyeglasses and moustaches, which could generate distracting flow patterns. Wrongly located face centers were however autonomously rejected by the model-based method, which could successfully assist to locate the face center in the remaining 8% of the sequences. All side regions were located correctly.

VII. CONCLUSION

Evaluating the trajectory of several face parts using the optical flow of lines is the main novelty of the proposed system. The liveness detection is successful in separating live face sequences from imitated photographs in motion, with an error below 1% on the test data. Although restricted to line velocity estimation, the suggested OFL is able to deliver robust measurements for face liveness assessment. A quick method for face center detection by optical flow pattern matching is shown feasible as well. The face part detection by model-based Gabor feature classification is robust to typical sources of errors like glasses and facial hair, and it successfully monitors and reliably assists the previous method. The scheme was evaluated on the XM2VTS database having scale variations up to 10%.

REFERENCES

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

[2] E. S. Bigun, J. Bigun, B. Duc, and S. Fischer. Expert conciliation for multi modal person authentication systems by bayesian statistics. In J. Bigun, G. Chollet, and G. Borgefors, editors, *Audio and Video based Person Authentication - AVBPA97*, pages 291–300. Springer, 1997.

[3] J. Bigun. Speed, frequency, and orientation tuned 3-d gabor filter banks and their design. In *Proc. International Conference on Pattern Recognition, ICPR, Jerusalem*, pages C–184–187. IEEE Computer Society, 1994.

[4] J. Bigun, H. Fronthaler, and K. Kollreider. Assuring liveness in biometric identity authentication by real-time face tracking. In *CIHSPS2004 - IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety, Venice, Italy, 21-22 July*, pages 104–112. IEEE Catalog No. 04EX815, ISBN 0-7803-8381-8, 2004.

[5] J. Bigun and G. H. Granlund. Optimal orientation detection of linear symmetry. In *First International Conference on Computer Vision, ICCV, June 8–11, London*, pages 433–438. IEEE Computer Society Press, Washington, DC., 1987.

[6] J. Bigun, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE-PAMI*, 13(8):775–790, 1991.

[7] T. Choudhury, B. Clarkson, T. Jebara, and A. Pentland. Multimodal person recognition using unconstrained audio and video. In *2nd International Conference on Audio-Visual Biometric Person Authentication, Washington D.C*, 22–23 March 1999.

[8] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995.

[9] B. Duc, S. Fischer, and J. Bigun. Face authentication with Gabor information on deformable graphs. *IEEE Trans. on Image Processing*, 8(4):504–516, 1999.

[10] Ian R Fasel, M. S Bartlett, and J. R. Movellan. A comparison of gabor methods for automatic detection of facial landmarks. In *International conference on Automatic Face and Gesture Recognition*, pages 242–248, May 2002.

[11] D. Gabor. Theory of communication. *Journal of the IEE*, 93:429–457, 1946.

[12] G. H. Granlund. In search of a general picture processing operator. *Computer Graphics and Image Processing*, 8(2):155–173, October 1978.

[13] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *J. physiology (London)*, 148, 1959.

[14] A. Jain, L. Hong, and Y. Kulkarni. A multimodal biometric system using fingerprint, face and speech. In *Audio and Video based Person Authentication - AVBPA99*, pages 182–187, 1999.

[15] H. Knutsson. *Filtering and reconstruction in image processing*. PhD Thesis no:88, Linköping University, ISY Bildbehandling S-581 83 Linkoeping, 1982.

[16] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Hurtz, and W. Konen. Distortion invariant object recognition in the dynamic link architectures. *IEEE Trans. on Computers*, 42(3):300–311, March 1993.

[17] J. Li, Y. Wang, T. Tan, and A. K. Jain. Live face detection based on the analysis of fourier spectra. In *Biometric Technology for Human Identification*, pages 296–303. SPIE Volume: 5404, August 2004.

[18] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. Xm2vtsdb: The extended m2vts database. In *Audio and Video based Person Authentication - AVBPA99*, pages 72–77. University of Maryland, 1999.

[19] G. A. Orban. *Neuronal operations in the visual cortex.* studies of brain functions. Springer, 1984.

[20] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40(2):614–634, 2001.

[21] Stephanie A. C. Schuckers. Spoofing and anti-spoofing measures. *Information Security Technical Report*, 7(4):56–62, 2002.

[22] F. Smeraldi and J. Bigun. Facial features detection by saccadic exploration of the Gabor decomposition. In *International Conference on Image Processing, ICIP-98, Chicago, October 4-7*, volume 3, pages 163–167, 1998.

[23] F. Smeraldi and J. Bigun. Retinal vision applied to facial features detection and face authentication. *Pattern Recognition Letters*, 23:463–475, 2002.

[24] F. Smeraldi, O. Carmona, and J. Bigun. Real–time head tracking by saccadic exploration and gabor decomposition. In A. T. Almeida and H. Araujo, editors, *Proc. the 5th International Workshop on Advanced Motion Control*, volume IEEE Cat. Num. 98TH8354, pages 684–687. IEEE Service Center, 445 Hoes Lane, P.O. Box 1331 Piscataway, NJ 08855-1331, USA, 1998.