EURASIP Journal on Image and Video Processing

SpringerOpen

Face liveness detection using dynamic texture

EURASIP Journal on Image and Video Processing 2014, **2014**:2 doi:10.1186/1687-5281-2014-2

Tiago de Freitas Pereira (tiagofrepereira@gmail.com)
Jukka Komulainen (jukmatt@ee.oulu.fi)
André Anjos (andre.anjos@idiap.ch)
José Mario De Martino (martino@fee.unicamp.br)
Abdenour Hadid (hadid@ee.oulu.fi)
Matti Pietikäinen (mkp@ee.oulu.fi)
Sébastien Marcel (marcel@idiap.ch)

ISSN	1687-5281
Article type	Research
Submission date	30 November 2012
Acceptance date	20 November 2013
Publication date	7 January 2014
Article URL	http://jivp.eurasipjournals.com/content/2014/1/2

Face liveness detection using dynamic texture

Tiago de Freitas Pereira[1*]
*Corresponding author
Email: tiagofrepereira@gmail.com

Jukka Komulainen[2]
Email: jukmatt@ee.oulu.fi

André Anjos[4]
Email: andre.anjos@idiap.ch

José Mario De Martino[3]
Email: martino@fee.unicamp.br

Abdenour Hadid[2]
Email: hadid@ee.oulu.fi

Matti Pietikäinen[2]
Email: mkp@ee.oulu.fi

Sébastien Marcel[4]
Email: marcel@idiap.ch

[1]CPqD Telecom & IT Solutions, School of Electrical and Computer Engineering, University of Campinas (UNICAMP), Campinas, São Paulo 13083-970, Brazil

[2]Center for Machine Vision Research, Department of Computer Science and Engineering, University of Oulu, Oulu FI-90014 , Finland

[3]School of Electrical and Computer Engineering, University of Campinas (UNICAMP), Campinas, São Paulo 13083-970, Brazil

[4]IDIAP Research Institute, Martigny CH-1920, Switzerland

Abstract

User authentication is an important step to protect information, and in this context, face biometrics is potentially advantageous. Face biometrics is natural, intuitive, easy to use, and less human-invasive. Unfortunately, recent work has revealed that face biometrics is vulnerable to spoofing attacks using cheap low-tech equipment. This paper introduces a novel and appealing approach to detect face spoofing using the spatiotemporal (dynamic texture) extensions of the highly popular local binary pattern operator. The key idea of the approach is to learn and detect the structure and the dynamics of the facial micro-textures that characterise real faces but not fake ones. We evaluated the approach with two publicly available databases (Replay-Attack Database and CASIA Face Anti-Spoofing Database). The results show that our approach performs better than state-of-the-art techniques following the provided evaluation protocols of each database.

Keywords

Anti-spoofing; Liveness detection; Countermeasure; Face recognition; Biometrics

1 Introduction

Because of its natural and non-intrusive interaction, identity verification and recognition using facial information are among the most active and challenging areas in computer vision research. Despite the significant progress of face recognition technology in the recent decades, a wide range of viewpoints, ageing of subjects and complex outdoor lighting are still research challenges. Advances in the area were extensively reported in [1] and [2].

Unfortunately, the issue of verifying if the face presented to a camera is indeed a face from a real person and not an attempt to deceive (spoof) the system has mostly been overlooked. It was not until very recently that the problem of spoofing attacks against face biometric system gained attention of the research community. This can be attested by the gradually increasing number of publicly available databases [3-6] and the recently organized IJCB 2011 competition on countermeasures to 2-D facial spoofing attacks [7] which was the first competition conducted for studying best practices for non-intrusive spoofing detection.

A spoofing attack consists in the use of forged biometric traits to gain illegitimate access to secured resources protected by a biometric authentication system. The lack of resistance to direct attacks is not exclusive to face biometrics. The findings in [8], [9] and [10] indicate that fingerprint authentication systems suffer from a similar weakness. The same shortcoming on iris recognition systems has been diagnosed [11-13]. Finally, in [14] and [15], the spoofing attacks to speaker biometrics are addressed. The literature review for spoofing in face recognition systems will be presented in Section 2.

In authentication systems based on face biometrics, spoofing attacks are usually perpetrated using photographs, videos or forged masks. While one can also use make-up or plastic surgery as means of spoofing, photographs and videos are probably the most common sources of spoofing attacks. Moreover, due to the increasing popularity of social network websites (Facebook, Flickr, YouTube, Instagram and others), a great deal of multimedia content - especially videos and photographs - is available on the web that can be used to spoof a face authentication system. In order to mitigate the vulnerability of face authentication systems, effective countermeasures against face spoofing have to be deployed.

Micro-texture analysis has been effectively used in detecting photo attacks from single face images [3,16,17]. Recently, the micro-texture-based analysis for spoofing detection was extended in the spatiotemporal domain in [18] and [19]. In both papers, the authors introduced a compact face liveness description that combines facial appearance and dynamics using spatiotemporal (dynamic texture) extensions of the highly popular local binary pattern (LBP) approach [20]. More specifically, local binary patterns from three orthogonal planes (LBP-TOP) were considered. This variant has shown to be very effective in describing the horizontal and vertical motion patterns in addition to appearance [21].

Even though authors of [18] and [19] considered LBP-TOP-based dynamic texture analysis for face spoofing detection, very dissimilar strategies were introduced for exploring the temporal dimension. In [18], the LBP-TOP-based face liveness description was extracted from relatively short time windows using the dense sampling of multiresolution approach, whereas an average of LBP-TOP features over longer temporal windows was used in [19]. Moreover, the experimental setups had significant differences because different face normalization techniques were applied in each work. Furthermore, the evaluations were performed on different databases (Replay-Attack Database [3] and CASIA Face Anti-Spoofing Database [6], respectively). In this article, we consolidate the methods proposed in [18] and [19], isolating the different variables and studying the potential of the different LBP-TOP countermeasures in different settings on both datasets. Furthermore, we demonstrate that our principled approach is able to consistently outperform prior work on the same databases and following the same evaluation protocols. We also provide an open-source framework that makes our research fully reproducible with minimal effort.

This work provides an in-depth analysis on the use of dynamic texture for face liveness description. We apply a unified experimental setup and evaluation methodology for assessing the effectiveness of the different temporal processing strategies introduced in [18] and [19]. The remainder of the paper is organized as follows: in Section 2, a brief review of the relevant literature is provided. The basic theory of local binary patterns in spatiotemporal domain is introduced in Section 3. Our dynamic texture-based face liveness description is described in Section 4. Section 5 presents the two publicly available databases which are used for evaluating the proposed countermeasure. In Section 6, we report on the experimental setup and results. Finally, in Section 7, we summarize this work highlighting its main contributions.

2 Literature review

Considering the type of countermeasures for face anti-spoofing that does not require user collaboration, Chakka et al. in [7] propose a classification scheme based on the following cues:

- Presence of vitality (liveness)
- Differences in motion patterns
- Differences in image quality assessment

Presence of vitality or liveness detection consists of search for features that only live faces can possess. For instance, Pan et al. in [4] exploited the observation that humans blink once every 2 to 4 s and proposed an eye blink-based countermeasure. Experiments carried out with the ZJU Eye Blink Database (http://www.cs.zju.edu.cn/gpan/database/db_blink.html) showed an accuracy of 95.7%.

The countermeasures based on differences in motion patterns rely on the fact that real faces display a different motion behaviour compared to a spoof attempt. Kollreider et al. [22] present a motion-based countermeasure that estimates the correlation between different regions of the face using optical flow field. In this approach, the input is considered a spoof if the optical flow field on the center of the face and on the center of the ears present the same direction. The performance was evaluated using the subset 'Head Rotation Shot' of the XM2VTS database whose real access was the videos of this subset, and the attacks were generated with hard copies of those data. Using this database, which was not made publicly available, an equal error rate (EER) of 0.5% was achieved. Anjos and Marcel [23] present a motion-based countermeasure measuring the correlation between the face and the background through simple frame differences. Using the PRINT ATTACK database, that approach presented a good discrimination power (half total error rate (HTER) equals to 9%).

Countermeasures based on differences in image quality assessment rely on the presence of artefacts intrinsically present at the attack media. Such remarkable properties can be originated from media quality issues or differences in reflectance properties of the object exposed to the camera. Li et al. [24] hypothesize that fraudulent photographs have less high-frequency components than real ones. To test the hypothesis, a small database was built with four identities containing both real access and printed photo attacks. With this private database, an accuracy of 100% was achieved. Assuming that real access images concentrate more information in a specific frequency band, Tan et al. [5] and Zhang et al. [6] used, as countermeasure, a set of difference of Gaussian filters (DoG) to select a specific frequency band to discriminate attacks and non-attacks. Evaluations carried out with the CASIA Face Anti-Spoofing Database and NUAA Photograph Imposter Database (http://parnec.nuaa.edu.cn/xtan/data/NUAAImposterDB.html) showed an equal error rate of 17% and an accuracy of 86%, respectively.

Because of differences in reflectance properties, real faces very likely present different texture patterns compared with fake faces. Following that hypothesis, Määttä et al. [17] and Chingovska et al. [3] explored the power of local binary patterns (LBP) as a countermeasure. Määttä et al. combined three different LBP configurations ($\text{LBP}^{u2}_{8,2}$, $\text{LBP}^{u2}_{16,2}$ and $\text{LBP}^{u2}_{8,1}$) in a normalized face image and trained a support vector machine (SVM) classifier to discriminate real and fake faces. Evaluations carried out with NUAA Photograph Impostor Database [5] showed a good discrimination power (2.9% in EER). Chingovska et al. analysed the effectiveness of $\text{LBP}^{u2}_{8,1}$ and set of extended LBPs [25] in still images to discriminate real and fake faces. Evaluations carried out with three different databases, the NUAA Photograph Impostor Database, Replay-Attack database and CASIA Face Anti-Spoofing Database [6], showed a good discrimination power with a HTER equal to 15.16%, 19.03% and 18.17%, respectively.

3 LBP-based dynamic texture description

Määttä et al. [17] and Chingovska et al. [3] propose a LBP-based countermeasures to spoofing attacks based on the hypothesis that real faces present different texture patterns in comparison with fake ones. However, the proposed techniques analyse each frame in isolation, not considering the behaviour over time. As pointed out in Section 2, motion is a cue explored in some works and in combination with texture can generate a powerful countermeasure. For describing the face liveness for spoofing detection, we considered a spatiotemporal representation which combines facial appearance and dynamics. We adopted the LBP-based spatiotemporal representation because of its recent convincing performance in modelling moving faces and facial expression recognition and also for dynamic texture recognition [20].

The LBP texture analysis operator, introduced by Ojala et al. [26,27], is defined as a gray-scale invariant texture measure, derived from a general definition of texture in a local neighbourhood. It is a powerful texture descriptor, and among its properties in real-world applications are its discriminative power, computational simplicity and tolerance against monotonic gray-scale changes. The original LBP operator forms labels for the image pixels by thresholding the 3×3 neighbourhood with the center value and considering the result as a binary number. The histogram of these $2^8 = 256$ different labels is then used as an image descriptor.

The original LBP operator was defined to only deal with the spatial information. However, more recently, it has been extended to a spatiotemporal representation for dynamic texture (DT) analysis. This has yielded to the so-called volume local binary pattern operator (VLBP) [21]. The idea behind VLBP consists of looking at dynamic texture (video sequence) as a set of volumes in the (X, Y, T) space where X and Y denote the spatial coordinates and T denotes the frame index (time). The neighborhood of each pixel is thus defined in a three-dimensional space. Then, similar to basic LBP in spatial domain, volume textons can be defined and extracted into histograms. Therefore, VLBP combines motion and appearance into a dynamic texture description.

To make VLBP computationally treatable and easy to extend, the co-occurrences of the LBP on the three orthogonal planes (LBP-TOP) was also introduced [21]. LBP-TOP consists of the three orthogonal planes - XY, XT and YT - and the concatenation of local binary pattern co-occurrence statistics in these three directions. The circular neighbourhoods are generalized to elliptical sampling to fit to the space-time statistics. The LBP codes are extracted from the XY, XT and YT planes, which are denoted as XY-LBP, XT-LBP and YT-LBP, for all pixels, and statistics of the three different planes are obtained and concatenated into a single histogram. The procedure is shown in Figure 1. In this representation, DT is encoded by the XY-LBP, XT-LBP and YT-LBP.

Figure 1 LBP from three orthogonal planes. **(a)** Three planes intersecting one pixel. **(b)** LBP histogram of each plane. **(c)** Concatenating the histograms (courtesy of [21]).

Using equal radii for the time and spatial axes is not a good choice for dynamic textures [21], and therefore, in the XT and YT planes, different radii can be assigned to sample neighbouring points in space and time. More generally, the radii R_x, R_x and R_t, respectively, in axes X, Y and T and the number of neighbouring points P_{XY}, P_{XT} and P_{YT}, respectively, in the XY, XT and YT planes can also be different. Furthermore, the type of LBP operator on each plane can vary; for example, the uniform pattern ($u2$) or rotation invariant uniform pattern ($riu2$) variants [20] can be deployed. The corresponding feature is denoted as LBP-TOP$^{\text{operator}}_{P_{XY},P_{XT},P_{YT},R_x,R_y,R_t}$.

Assuming we are given a $X \times Y \times T$ dynamic texture $(x_c \in \{0, \cdots, X-1\}, y_c \in \{0, \cdots, Y-1\}, t_c \in \{0, \cdots, T-1\})$, i.e. a video sequence. A histogram of the DT can be defined as

$$H_{i,j} = \sum_{x,y,t} I\{f_j(x,y,t) = i\}, \quad i = 0, \cdots, n_j - 1; j = 0, 1, 2 \tag{1}$$

where n_j is the number of different labels produced by the LBP operator in the jth plane ($j = 0 : XY,\ 1 : XT\ and\ 2 : YT$), and $f_i(x,y,t)$ expresses the LBP code of the central pixel (x,y,t) in the jth plane.

Similar to the original LBP, the histograms must be normalized to get a coherent description for comparing the DTs:

$$N_{i,j} = \frac{H_{i,j}}{\sum_{k=0}^{n_j-1} H_{k,j}} \ . \tag{2}$$

In addition to the computational simplification, compared with VLBP, LBP-TOP has the advantage to generate independent histograms for each of the intersecting planes, in space and time, which can be treated in combination or individually. Because of the aforementioned complexity issues on the implementation of a VLBP-based processor, the developed spatiotemporal face liveness description uses LBP-TOP to encode both facial appearance and dynamics.

Our key idea is to learn and detect the structure and the dynamics of the facial micro-textures that characterise real faces but not fake ones. Due to its tolerance against monotonic gray-scale changes, LBP-based representation is adequate for measuring the facial texture quality and determining whether degradations due to recapturing process, e.g. the used spoofing medium, are observed. Instead of just applying static texture analysis, we exploit also several dynamic visual cues that are based on either the motion patterns of a genuine human face or the used display medium.

Unlike photographs and display devices, real faces are indeed non-rigid objects with contractions of facial muscles which result in temporally deformed facial features such as eye lids and lips. Therefore, it can be assumed that the specific facial motion patterns (including eye blinking, mouth movements and facial expression changes) should be detected when a live human being is observed in front of the camera. The movement of the display medium may cause several distinctive motion patterns that do not describe genuine faces. As shown in Figure 2, the use of (planar) spoofing medium might cause sudden characteristic reflections when a photograph is warped or because of a glossy surface of the display medium. As it can be seen, warped photo attacks may cause also distorted facial motion patterns. It is likely that hand-held attacks introduce synchronized shaking of the face and spoofing medium which can be observed as excessive relative motion in the view and facial region if the distance between the display medium and the camera is relatively short. In this work, we try to exploit the aforementioned visual cues for face spoofing detection by exploring the dynamic texture content of the facial region. We adopted the LBP-based spoofing detection in spatiotemporal domain because LBP-TOP features have been successfully applied in describing dynamic events, e.g. facial expressions [21].

Figure 2 Example sequence of a warped photo attack from the CASIA Face Anti-Spoofing Database [6]. This describes the characteristic reflections (flickering) of a planar spoofing medium and the distorted motion patterns.

4 The proposed countermeasure

Figure 3 shows a block diagram of the proposed countermeasure. First, each frame of the original frame sequence was gray-scaled and passed through a face detector using modified census transform (MCT) features [28]. Only detected faces with more than 50 pixels of width and height were considered. The detected faces were geometric normalized to 64×64 pixels. In order to reduce the face detector noise, the same face bounding box was used for each set of frames used in the LBP-TOP calculation. As can be seen in the Figure 4, the middle frame was chosen. Unfortunately, the face detector is not error free, and in case of error in the middle frame face detection, the nearest detection was chosen; otherwise, the observation was discarded. After the face detection step, the LBP operators were applied for each plane (XY, XT and YT) and the histograms were computed and then concatenated. After the feature extraction step, binary classification can be used to discriminate spoofing attacks from real access attempts.

Figure 3 Block diagram of the proposed countermeasure.

Figure 4 Face detection strategy for $R_t = 1$.

Face liveness is rather difficult to be determined based on the motion between a couple of successive frames. The used volume can be expanded along the temporal dimension by increasing R_t, as aforementioned in Section 3. This way to deal with dynamic texture is called single resolution approach, since only one histogram per LBP-TOP plane is accumulated. However, this leads to rather sparse sampling on the temporal planes XT and YT; thus, we might loose valuable details. In order to explore the dynamic texture information more carefully, we proposed the multiresolution approach.

The multiresolution approach can be performed by concatenating the histograms in the time domain (XT and YT) for different values of R_t. The notation chosen to represent these settings is using brackets for the multiresolution data. For example, $R_t = [1 - 3]$ means that the LBP-TOP operator will be calculated for $R_t = 1$, $R_t = 2$ and $R_t = 3$ and all resultant histograms will be concatenated. With the multiresolution approach, dense sampling on the temporal planes XT and YT is achieved.

The proposed countermeasure was implemented using the free signal processing and machine learning toolbox Bob [29], and the source code of the algorithm is available as an add-on package to this framework (http://pypi.python.org/pypi/antispoofing.lbptop). After installation, it is possible to reproduce all results reported in this article.

5 Spoofing databases

In this section, we give an overview of the two largest and most challenging face spoofing databases, Replay-Attack Database [3] and the CASIA Face Anti-Spoofing Database [6], consisting of real access attempts and several fake face attacks of different natures under varying conditions. Instead of still images, both datasets contain short video recordings which makes them suitable for evaluating countermeasures that exploit also temporal information.

5.1 Replay-Attack Database

The Replay-Attack Database (http://www.idiap.ch/dataset/replayattack) [3] consists of short video (∼10s) recordings of both real-access and attack attempts to 50 different identities using a laptop. It contains 1,200 videos (200 real-access and 1,000 attacks), and the attacks were taken in three different scenarios with two different illumination and support conditions. The scenarios of attack include the following:

1. *Print*: the attacker displays hard copies of high-resolution photographs printed on A4 paper
2. *Mobile*: the attacker displays photos and videos taken with an iPhone 3GS using the phone screen
3. *Highdef*: the attacker displays high-resolution photos and videos using an iPad screen with a resolution of $1,024 \times 768$.

The illumination conditions include the following:

1. *Controlled*: the background of the scene is uniform and the light of a fluorescent lamp illuminates the scene
2. *Adverse*: the background of the scene is non-uniform and daylight illuminates the scene

The support conditions include the following:

1. *Hand-based*: the attacker holds the attack media using his own hands
2. *Fixed*: the attacker sets the attack device in a fixed support so it does not move during the spoofing attempt

Figure 5 shows some examples of real accesses and attacks in different scenarios. The top row shows samples from the controlled scenario. The bottom row shows samples from the adverse scenario. Columns from left to right show examples of real access, printed photograph, mobile phone and tablet attacks.

Figure 5 Some frames of real access and spoofing attempts (courtesy of [3]).

The Replay-Attack Database provides a protocol for objectively evaluating a given countermeasure. Such protocol defines three non-overlapping partitions for training, development and testing countermeasures (see Table 1). The training set should be used to train the countermeasure, and the development set is used to tune the countermeasure and to estimate a threshold value to be used in the test set. The test set must be used only to report results. As a performance measurement, the protocol advises the use of HTER (Equation 3).

$$\mathrm{HTER} = \frac{\mathrm{FAR}(\tau, D) + \mathrm{FRR}(\tau, D)}{2}, \tag{3}$$

where τ is a threshold, D is the dataset, FAR is the false acceptance rate and FRR is the false rejection rate. In this protocol, the value of τ is estimated on the EER using the development set.

Table 1 Number of videos in each subset

Type	Train	Devel.	Test	Total
Real access	60	60	80	200
Print attack	$30 + 30$	$30 + 30$	$40 + 40$	$100 + 100$
Mobile attack	$60 + 60$	$60 + 60$	$80 + 80$	$200 + 200$
Highdef attack	$60 + 60$	$60 + 60$	$80 + 80$	$200 + 200$
Total	360	360	480	1200

Numbers displayed as sums indicate the amount of hand-based and fixed support attack available in each subset [3].

5.2 CASIA Face Anti-Spoofing Database

The CASIA Face Anti-Spoofing Database (http://www.cbsr.ia.ac.cn/english/FaceAntiSpoof%20Databases.asp) [6] contains 50 real clients, and the corresponding fake faces are captured with high quality from the original ones. The variety is achieved by introducing three imaging qualities (low, normal and high) and three fake face attacks which include warped photo, cut photo (eyeblink) and video attacks. Examples from the database can be seen in Figure 6. Altogether, the database consists of 600 video clips, and the subjects are divided into subsets for training and testing (240 and 360, respectively). Results of a baseline system are also provided along the database for fair comparison. The baseline system considers the high-frequency information in the facial region using multiple DoG features and SVM classifier and is inspired by the work of Tan et al. [5].

Figure 6 Example images of real accesses and the corresponding spoofing attempts (courtesy of [6]).

Since the main purpose of the database is to investigate the possible effects of different fake face types and imaging qualities, the test protocol consists of seven scenarios in which particular train and test samples are to be used. The quality test considers the three imaging qualities separately, low (1), normal (2) and high quality (3), and evaluates the overall spoofing detection performance under a variety of attacks at the given imaging quality. Similarly, the fake face test assesses how robust the anti-spoofing measure is to specific fake face attacks, warped photo (4), cut photo (5) and video attacks (6), regardless of the imaging quality. In the overall test (7), all data are used to give a more general evaluation. The results of each scenario are reported as detection error trade-off (DET) curves and EERs, which is the point where FAR equals FRR on the DET curve.

6 Experiments

This section provides an in-depth analysis on the proposed LBP-TOP-based face liveness description using the Replay-Attack Database [3] and the CASIA Face Anti-Spoofing Database [6]. First, we study the effect of different classifiers and LBP-TOP parameters by following the evaluation method proposed in [18]. The LBP-TOP representation is computed over relatively short temporal windows, and the results are reported using the overall classification accuracy for the individual volumes. Altogether, four experiments were carried out evaluating the effectiveness of

1. Each LBP-TOP plane individually and in combination
2. Different classifiers
3. Different LBP operators
4. The multiresolution approach

In order to study the effect of the different variables, each parameter was tuned solely (fixing other elements) using the development set of each face spoofing database. It should be noted that unlike the Replay-Attack Database, the CASIA Face Anti-Spoofing Database is lacking a specific development set. Therefore, the first 4 experiments were performed in this database using cross-validation by randomly dividing the training data into fivefold. Hence, the results presented for CASIA Face Anti-Spoofing Database are actually the average HTER on the test set over five iterations of the algorithm with different folds playing the role of a development set.

Finally, we also studied the accumulation of facial appearance and dynamics information over longer time windows and perform an evaluation at system level. The access attempt-based results presented in Section 6.5 were obtained using the official protocol of each database.

Inspired by [3], the LBP-TOP operator chosen to start the evaluation was LBP-TOP$^{u2}_{8,8,8,1,1,R_t}$.

6.1 Effectiveness of each LBP-TOP plane individually and in combination

In this experiment, we analysed the effectiveness of each individual plane and their combinations when the multiresolution area is increased. Figure 7 shows the HTER evolution, on the test set, considering individual and combined histograms of LBP-TOP planes for each database. We used, as binary classifier, a linear projection derived from linear discriminant analysis (LDA) as in [3].

Figure 7 Evaluation of HTER (%) in each plane when multiresolution area (R_t) is increased. With LBP-TOP$^{u2}_{8,8,8,1,1,R_t}$ and LDA classifier test set. **(a)** Replay-Attack Database. **(b)** CASIA Face Anti-Spoofing Database.

The results indicate differences in the performance between the two databases. The temporal components (XT and YT) are a decisive cue for the Replay-Attack Database, and the combination of all three planes (XY, XT and YT) gives the best performance. Conversely, for the CASIA Face Anti-Spoofing Database, the addition of temporal planes improves the performance only slightly compared to the spatial LBP representation (considering only the XY plane). These observations can be explained by taking a closer look at the differences in the databases and their spoofing attack scenarios. 2-D fake face attacks can be categorized into two groups, close-up and scenic attacks, based on how the fake face is represented with the spoofing medium.

A close-up spoof describes only the facial area which is presented to the sensor. The main weakness with the tightly cropped fake faces is that the boundaries of the spoofing medium, e.g. a video screen frame, photograph edges or the attacker's hands, are usually visible during the attack and thus can be detected in the scene [19]. However, these visual cues can be hidden by incorporating the background scene in the face spoof and placing the resulting scenic fake face very near to the sensor as performed on the Replay-Attack Database. In such cases, the description of facial appearance leads to rather good performance because the proximity between the spoofing medium and the camera causes the recaptured face image to be out-of-focus also revealing other facial texture quality issues, like degradation due to the used spoofing medium. Furthermore, the attacks in Replay-Attack Database are performed using two types of support conditions, fixed and hand-held. Naturally, the LBP-TOP-based face representation can easily detect fixed photo and print attacks since there is no variation in the facial texture over time. On the other hand, the hand-held attacks introduce synchronized shaking of the face and spoofing medium. This can be observed as excessive relative motion in the view, again, due to the proximity between the display medium and the sensor. Since the distinctive global motion patterns are clearly visible also on the facial region, they can be captured even by computing the LBP-TOP description over relatively short temporal windows, i.e. low values of R_t.

In contrast, the CASIA Face Anti-Spoofing Database consists of close-up face spoofs. The distance between the camera and the display medium is much farther compared to the attacks on Replay-Attack Database. The display medium does not usually move much in the attack scenarios. Therefore, the overall translational movement of a fake face is much closer to the motion of a genuine head. Due to the lack of distinctive shaking of the display medium, the CASIA Face Anti-Spoofing Database can be considered to be more challenging from the dynamic texture point of view. Because the motion cues are harder to explore in some attack scenarios using small values of R_t, we investigated in Section 6.5 whether the use of longer time windows helps to reveal the disparities between a genuine face and a fake one.

6.2 Effectiveness of different classifiers

In this experiment, we analysed the effectiveness of different classifiers when the multiresolution area is increased. Figure 8 shows the HTER evolution, on the test set, under three different classification schemes. The first one uses χ^2 distance, since the feature vectors are histograms. The same strategy reported in [3] was carried out. A reference histogram only with real accesses was created averaging the histograms in the training set. The last two selected classification schemes analysed were LDA and SVM with a radial basis function kernel (RBF).

Figure 8 Evaluation of HTER (%) with LBP-TOP$^{u2}_{8,8,8,1,1,R_t}$ using different classifiers. (a) Replay-Attack Database. **(b)** CASIA Face Anti-Spoofing Database.

The SVM classifier with an RBF kernel provided the best performance on the Replay-Attack Database and the CASIA Face Anti-Spoofing Database (7.97% and 20.72% in terms of HTER, respectively). However, it is important to remark that the same LBP-TOP configuration with an LDA classifier resulted in comparable performance (11.35% and 24.91% in terms of HTER). This is not a huge gap, and the classification scheme is far simpler. As similar findings have been reported [3,30], the use of simple and computationally efficient classifiers should be indeed considered when constructing real-world anti-spoofing solutions.

6.3 Effectiveness of different LBP operators

The size of the histogram in a multiresolution analysis, in time domain, increases linearly with R_t. The choice of an appropriate LBP representation in the planes is an important issue since it impacts the size of the histograms. Using uniform patterns or rotation invariant extensions, in one or multiple planes, may bring a significant reduction in computational complexity. In this experiment, the effectiveness of different LBP operators in the three LBP-TOP planes (XY, XT and YT) was analysed. Figure 9 shows the performance, in HTER terms, configuring each plane as basic LBP (with 256 bins for $P = 8$), LBP^{u2} (uniform patterns) and LBP^{riu2} (rotation invariant uniform patterns) when the multiresolution area (R_t) is increased in both databases. Results must be interpreted with the support of Figure 10, which shows the number of bins on the histograms used for classifications in each configuration.

Figure 9 Evaluation of HTER (%) with LBP-TOP$_{8,8,8,1,1,R_t}$ using different LBP configurations in planes with SVM classifier. **(a)** Replay-Attack Database **(b)** CASIA Face Anti-Spoofing Database.

Figure 10 Evaluation of the histogram size when (R_t) is increased.

When the multiresolution area is increased, the HTER saturates for LBP^{riu2} and LBP^{u2} on both datasets. For the basic LBP operator, a minimum can be observed in 7.60% and 20.71% on the Replay-Attack Database and CASIA Face Anti-Spoofing Database, respectively. On both databases, basic LBP and LBP^{u2} presented similar performance. Even though the use of regular LBP leads to the best results, the LBP^{u2} operator seems to provide a reasonable trade-off between computational complexity (see Figure 10) and performance. Hence, we will still proceed with LBP^{u2}.

6.4 Effectiveness of the multiresolution approach

In this experiment, we analysed the effectiveness of the multiresolution approach in comparison with the single resolution approach. The single resolution approach consists of using only fixed values for R_t, without concatenating histograms for each R_t. With this approach, the size of the histograms will be constant for different values of R_t, which decreases the computational complexity compared to the multiresolution approach. Figure 11 shows the HTER evolution for different values of R_t in both databases comparing both approaches.

Figure 11 Evaluation of HTER (%) using LBP-TOP$^{u2}_{8,8,8,1,1,R_t}$ with single resolution and multiresolution approach using SVM classifier. **(a)** Replay-Attack Database. **(b)** CASIA Face Anti-Spoofing database.

On both datasets, the HTER of the single resolution approach increases with R_t, whereas the multiresolution approach helps to keep the HTER low when the multiresolution area is increased. This suggests that the increase of R_t causes more sparse sampling in the single resolution approach when valuable motion information is lost. In contrary, the more dense sampling of the multiresolution approach is able to provide a more detailed description of the motion patterns, thus improving the discriminative power.

6.5 Access attempt-based analysis

In the previous experiments, the importance of the temporal dimension was studied using the single resolution and the multiresolution approaches. As seen in Section 6.1, the multiresolution approach is able to capture well the nature of fixed photo attacks and the excessive motion of display medium, especially on the Replay-Attack Database. However, in some attack scenarios, the motion patterns were harder to explore using small values of R_t. Therefore, we now study how the used temporal window size affects the performance when the facial appearance and dynamics information are accumulated over time. The face description of the single resolution and multiresolution methods can be accumulated over longer time periods either by averaging the features within a time window or by classifying each subvolume and then averaging the scores within the current window. In this manner, we are able to provide dense temporal sampling over longer temporal windows without excessively increasing the size of the feature histogram.

To follow the method used in previous experiments, we begin evaluating the two averaging strategies with the LBP-TOP$^{u2}_{8,8,8,1,1,1}$ operator and a SVM classifier with RBF kernel. In order to determine the video-based system performance, we applied both the average of features and scores on the first valid time window of N frames from the beginning of each video sequence. It should be noted that the following access attempt-based analysis is based on the official protocol of each database. Thus, the results on Replay-Attack Database are reported in terms of HTER, whereas the performance on CASIA Face Anti-Spoofing Database is described using EER.

The access attempt-based performance of both averaging strategies on the two databases is presented in Figure 12. The results indicate that when the amount of temporal information increases, the better we are able to discriminate real faces from fake ones. This is the case especially on the CASIA Face Anti-Spoofing Database in which the distinctive motion clues, such as the excessive shaking of the display medium, cannot be exploited. However, when longer video sequences are explored, we are more likely to observe other specific dynamic events, such as different facial motion patterns (including eye blinking, lip movements and facial expression changes) or sudden characteristic reflections of planar spoofing media which can be used for differentiating real faces from fake ones. It is also interesting to notice that by averaging features, more stable and robust spoofing detection performance is achieved on both databases. The averaging scores of individual subvolumes seem to suffer from outliers; thus, more sophisticated temporal processing of scores might lead to more stable behaviour.

Figure 12 Access attempt-based evaluation. Different time window sizes were evaluated using mean of features and mean of scores with LBP-TOP$^{u2}_{8,8,8,1,1,1}$. **(a)** Replay-Attack Database (HTER %). **(b)** CASIA Face Anti-Spoofing Database (EER %).

According to the official test protocol of CASIA Face Anti-Spoofing, also the DET curves and the EERs for the seven scenarios should be reported. Based on the previous analysis, we chose to use the average of features within a time window of 75 frames which corresponds to 3 s of video time. As it can be seen in Figure 13 and Table 2, the use of only facial appearance (LBP) leads to better results compared to the baseline method (CASIA baseline). More importantly, when the temporal planes XT and YT are also considered for spatiotemporal face description (LBP-TOP), a significant performance enhancement is obtained (from 16% to 10% in terms of EER), thus confirming the benefits of encoding and exploiting not only the facial appearance but also the facial dynamics information.

Figure 13 Overall test protocol on the CASIA Face Anti-Spoofing Database. Overall performance of LBP-TOP$^{u2}_{8,8,8,1,1,1}$ using the average of features compared to the DoG baseline method and LBP$^{u2}_{8,1}$.

Table 2 Comparison of EER (%)

Scenario	Low	Normal	High	Warped	Cut	Video	Overall
DoG baseline [6]	13	13	26	16	6	24	17
LBP$^{u2}_{8,1}$	11	17	13	13	16	16	16
LBP-TOP$^{u2}_{8,8,8,1,1,1}$	10	12	13	6	12	10	10

This table shows comparison between the DoG baseline method, LBP$^{u2}_{8,1}$ and LBP-TOP$^{u2}_{8,8,8,1,1,1}$ using the average of features on the CASIA Face Anti-Spoofing Database.

More detailed results for each scenario are presented in Figure 14 and in Table 2. The results indicate that the proposed LBP-TOP-based face description yields best results in all configurations except under cut-photo attacks. As described in [6], the DoG filtering baseline method is able to capture the less variational nature of the cut eye regions well. However, the difference in the motion patterns seems to be too small for our LBP-TOP-based approach as mainly eye blinking occurs during the cut-photo attacks and no other motion is present. The EER development presented in Table 3 supports this conclusion since the performance under cut-photo attacks does not improve that much if longer temporal window is applied compared to the other scenarios.

Figure 14 The different test protocols of the CASIA Face Anti-Spoofing Database. Performance of LBP-TOP$^{u2}_{8,8,8,1,1,1}$ using the average of features compared to the DoG baseline method and LBP$^{u2}_{8,1}$.

Table 3 Effect of different time window sizes on CASIA Face Anti-Spoofing Database

Frames	Low	Normal	High	Warped	Cut	Video
1	17	27	23	29	16	20
5	13	20	20	19	14	14
10	14	20	19	18	16	14
25	13	13	10	10	14	12
50	13	11	10	7	13	10
75	10	12	13	6	12	10

This table shows EER development of LBP-TOP$^{u2}_{8,8,8,1,1,1}$ using the average of features.

On the other hand, the spatiotemporal face description is able to improve the major drawbacks of DoG-based countermeasure. Unlike the baseline method, our approach performs almost equally well at all three imaging qualities. Furthermore, the performance under warped photo and video attacks is significantly better. Especially the characteristic specular reflections (flickering) and excessive and distorted motion of warped photo attacks can be described very well.

6.6 Summary

Tables 4 and 5 summarize all the results obtained for each database following their provided protocols. In order to be comparable with still frame analysis presented for example in [3], the results for the Replay-Attack Database represent the overall classification accuracy considering each frame individually. The access attempt-based results are reported only for the CASIA Face Anti-Spoofing Database as requested in its test protocol.

Table 4 HTER (%) of the best results on the Replay-Attack Database

	Dev	Test
Motion Correlation [23]	11.78	11.79
LBP$^{u2}_{8,1}$ + SVM	14.84	15.16
LBP$_{3\times3}$ + SVM [3]	13.90	13.87
LBP-TOP$^{u2}_{8,8,8,1,1,1}$ + SVM	8.17	8.51
LBP-TOP$_{8,8,8,1,1,[1-2]}$ + SVM	7.88	7.60

This table shows the HTER of the best results achieved on the Replay-Attack Database (following the database protocol) compared with the provided baseline.

Table 5 EER (%) of the best results on the CASIA Face Anti-Spoofing Database

	Test
DoG baseline [6]	17
LBP$^{u2}_{8,1}$ + SVM	16
LBP-TOP$^{u2}_{8,8,8,1,1,1}$ with average of features + SVM	10

This table shows the EER of the best results achieved on the CASIA Face Anti-Spoofing Database (following the database protocol) compared with the provided baseline.

Table 4 shows also the results for the LBP (http://pypi.python.org/pypi/antispoofing.lbp) [3] and the Motion Correlation (http://pypi.python.org/pypi/antispoofing.motion) [23] based countermeasures whose source code is freely available. Table 5 contains the provided DoG-based baseline and the holistic

LBP-based face description. It can be seen that the proposed countermeasure presented the best results, overtaking the baseline results in both databases, thus confirming the benefits of encoding and exploiting not only the facial appearance but also the facial dynamics information. Unfortunately, our comparison is limited to these countermeasures due to the lack of publicly available implementations of other state-of-the-art techniques presented in the literature.

During these experiments, we observed that the general performance of the proposed countermeasure was consistently better on the Replay-Attack Database compared to the CASIA Face Anti-Spoofing Database. As mentioned in Section 6.1, the nature of the attack scenarios is different between the two datasets. In the Replay-Attack Database, our LBP-TOP-based face description was able to capture motion patterns of fixed photo attacks and scenic fake face attacks already when only relatively short time windows were explored. Performances below 10% (HTER) were achieved. On the other hand, the CASIA Face Anti-Spoofing Database turned out to be more challenging from the dynamic texture point of view. Due to the lack of motion, analysis of longer temporal windows was required in order to find out distinctive motion patterns between genuine faces and fake ones. As it can be seen in Table 5, by extending the micro-texture-based spoofing detection into the spatiotemporal domain, an improvement from 16% to 10% in terms of EER was obtained. The results also indicate that the proposed dynamic texture-based face liveness description was able to improve the state of the art on both datasets.

7 Conclusion

Inspired by the recent progress in dynamic texture, the problem of face spoofing detection was recently investigated in two independent articles using spatiotemporal local binary patterns. The key idea of the proposed countermeasures consists of analysing the structure and the dynamics of the micro-textures in the facial regions using LBP-TOP features that provide an efficient and compact representation for face liveness description. However, very dissimilar strategies were introduced for exploring the temporal dimension even though the same features were utilized. Furthermore, the experiments were carried out using different face normalization techniques and different databases. In this article, we consolidated the methods proposed in the previous studies, isolating the different variables and studying the potential of the different LBP-TOP countermeasures in different settings on the two publicly available datasets. Furthermore, we also provided an open-source framework that makes our research fully reproducible with minimal effort.

Experiments carried out with a unified experimental setup and evaluation methodology showed that the dynamic texture-based countermeasure was able to consistently outperform prior work on both datasets. Best results were achieved using a nonlinear SVM classifier, but it is important to note that experiments with a simpler LDA-based classification scheme resulted in comparable performance under various spoofing attack scenarios. Thus, the use of simple and computationally efficient classifiers should be indeed considered when constructing real-world anti-spoofing solutions. In a future work, we will study the generalization capabilities of the proposed countermeasure using multiple face anti-spoofing databases. In other words, we plan to perform cross-database experiments by training and tuning the LBP-TOP-based face description solely on one dataset and test on another one.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work has been performed within the context of the TABULA RASA project, part of the 7th Framework Research Programme of the European Union (EU), under the grant agreement number 257289. The financial support of FUNTTEL (Brazilian Telecommunication Technological Development Fund), Academy of Finland and Infotech Oulu Doctoral Program is also gratefully acknowledge.

References

1. P Flynn, A Jain, A Ross, *Handbook of Biometrics*. (Springer, 2008)
2. S Li, A Jain, *Handbook of Face Recognition.* (Springer, 2011)
3. I Chingovska, A Anjos, S Marcel, On the effectiveness of local binary patterns in face anti-spoofing. in *IEEE International Conference of the Biometrics Special Interest Group*, Darmstadt, 6–7 September 2012
4. G Pan, L Sun, Z Wu, S Lao, Eyeblink-based anti-spoofing in face recognition from a generic webcamera, in *IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, 14–21 October 2007, pp. 1–8
5. X Tan, Y Li, J Liu, L Jiang, Face liveness detection from a single image with sparse low rank bilinear discriminative model, in *11th European Conference on Computer Vision: Part VI. ECCV'10,* Heraklion, Crete, Greece, 5–11 September 2010, pp. 504–517
6. Z Zhang, J Yan, S Liu, Z Lei, D Yi, SZ Li, A face antispoofing database with diverse attacks, in *Proceedings of 5th IAPR International Conference on Biometrics (ICB'12)*, New Delhi, India, 29 March - 1 April 2012
7. 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, ZS Li, WR Schwartz, A Rocha, H Pedrini, LJ Navarro, C-M Santana, J Määttä, A Hadid, M Pietikäinen, Competition on counter measures to 2-D facial spoofing attacks, in *IAPR IEEE International Joint Conference on Biometrics*, Washington DC, USA, 11–13 October 2011
8. U Uludag, A Jain, Attacks on biometric systems: a case study in fingerprints, in *Proc. SPIE-EI* San Rose CA, USA, 18–22 January ,pp. 622–633
9. J Leyden, Gummi bears defeat fingerprint sensors. The Register **16**, (2002)
10. T Matsumoto, H Matsumoto, K Yamada, S Hoshino, Impact of artificial gummy fingers on fingerprint systems, in *Proceedings of SPIE, Volume 4677,* San Jose CA, USA 24–25 January 2002, pp. 275–289
11. P Johnson, B Tan, S Schuckers, Multimodal fusion vulnerability to non-zero effort (spoof) imposters, in *IEEE Informational Workshop on Information Forensics and Security*, Seattle, USA, 12–15 December 2010, pp. 1–5
12. M Kanematsu, H Takano, K Nakamura, Highly reliable liveness detection method for iris recognition, in, *International Conference on Instrumentation, Control and Information Technology*, Takamatsu, 17–20 September 2007, pp. 361–364
13. A Pacut, A Czajka, A liveness detection for iris biometrics, in *40th Annual IEEE International Carnahan Conferences Security Technology*, Lexington, KY, October 2006, pp. 122–129

14. G Chetty, M Wagner, Liveness verification in audio-video speaker authentication, in *Proceeding of International Conference on Spoken Language Processing ICSLP, Volume 4* Jeju Island, Korea, 4–8 October 2004, pp. 2509–2512

15. N Eveno, L Besacier, A speaker independent“liveness” test for audio-visual biometrics, in *9th European Conference on Speech Communication and Technology*, Lisbon, 4–8 September 2005

16. J Bai, TT Ng, X Gao, YQ Shi, Is physics-based liveness detection truly possible with a single image?, in *IEEE International Symposium on Circuits and Systems (ISCAS)*, Paris, 30 May - 2 June 2010, pp. 3425–3428

17. J Määttä, A Hadid, M Pietikäinen, Face spoofing detection from single images using micro-texture analysis, in *IAPR IEEE International Joint Conference on Biometrics*, Washington DC, USA, 11–13 October 2011

18. TF Pereira, A Anjos, JM De Martino, S Marcel, LBP-TOP based countermeasure against facial spoofing attacks, in *International Workshop on Computer Vision With Local Binary Pattern Variants - ACCV)*, Daejeon, Korea, 5–6 November 2012

19. J Komulainen, A Hadid, M Pietikäinen, Face spoofing detection using dynamic texture, in *International Workshop on Computer Vision With Local Binary Pattern Variants - ACCV*, Daejeon, Korea, 5–6 November 2012

20. M Pietikäinen, A Hadid, G Zhao, T Ahonen, *Computer Vision Using Local Binary Patterns, Volume 40*. (Springer, 2011)

21. G Zhao, M Pietikäinen, Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans. Pattern Anal. Mach. Intell. **29**, 915–928 (2007)

22. K Kollreider, H Fronthaler, J Bigun, Non-intrusive liveness detection by face images. Elsevier Image and Vision Computing **27**, 233–244 (2009)

23. A Anjos, S Marcel, Counter-measures to photo attacks in face recognition: a public database and a baseline, in *IAPR IEEE International Joint Conference on Biometrics*, Washington DC, USA, 11–13 October 2011)

24. J Li, Y Wang, T Tan, A Jain, Live face detection based on the analysis of fourier spectra. Biometric Technology for Human Identification **5404**, 296–303 (2004)

25. J Trefnỳ, J Matas, Extended set of local binary patterns for rapid object detection, in *15th Computer Vision Winter Workshop, Volume 2010*, Czech Republic, 3–5 February 2010

26. T Ojala, M Pietikäinen, D Harwood, A comparative study of texture measures with classification based on feature distributions. Pattern Recognit. **29**, 51–59 (1996)

27. T Ojala, M Pietikäinen, T Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. on PAMI **24**, (2002)

28. B Froba, A Ernst, Face detection with the modified census transform, in, *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, Seoul, South Korea, 17–19 May 2004, pp. 91–96

29. A Anjos, L El Shafey, R Wallace, M Günther, C McCool, S Marcel, Bob: a free signal processing and machine learning toolbox for researchers, in *20th ACM Conference on Multimedia Systems*, Nara, Japan, 22–24 February 2012)

30. J Komulainen, A Anjos, A Hadid, S Marcel, M Pietikäinen, Complementary countermeasures for detecting scenic face spoofing attacks, in *6th IAPR International Conference on Biometrics*, Madrid, 4–7 June 2013

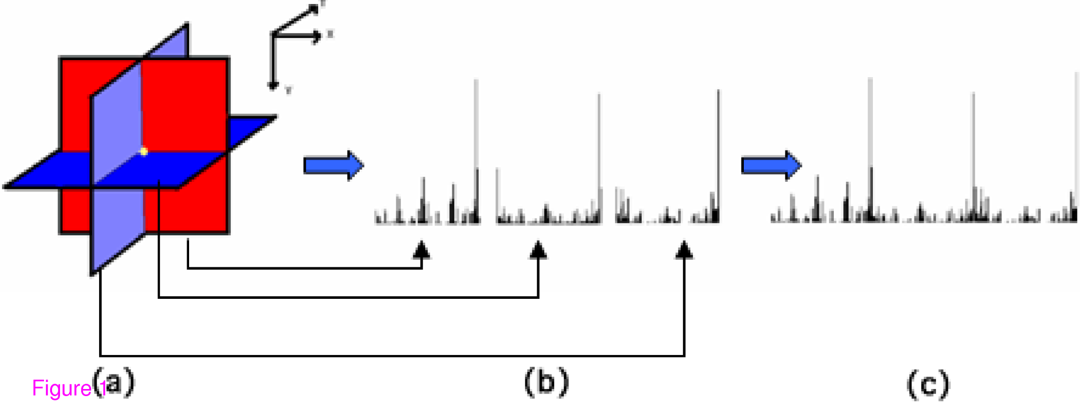


Figure 1

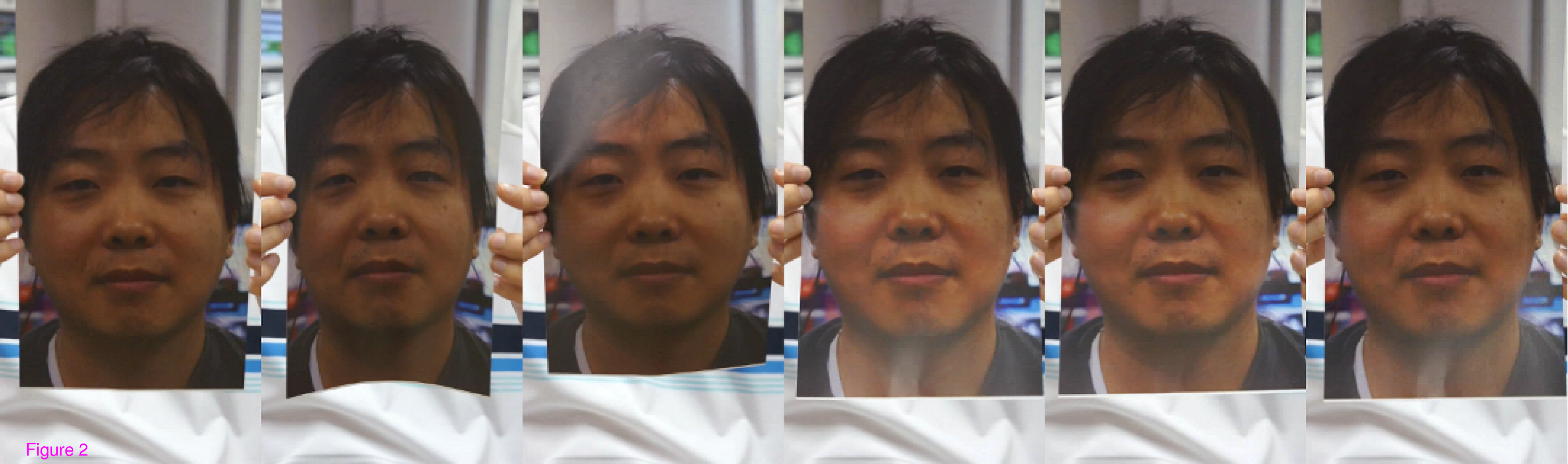


Figure 2

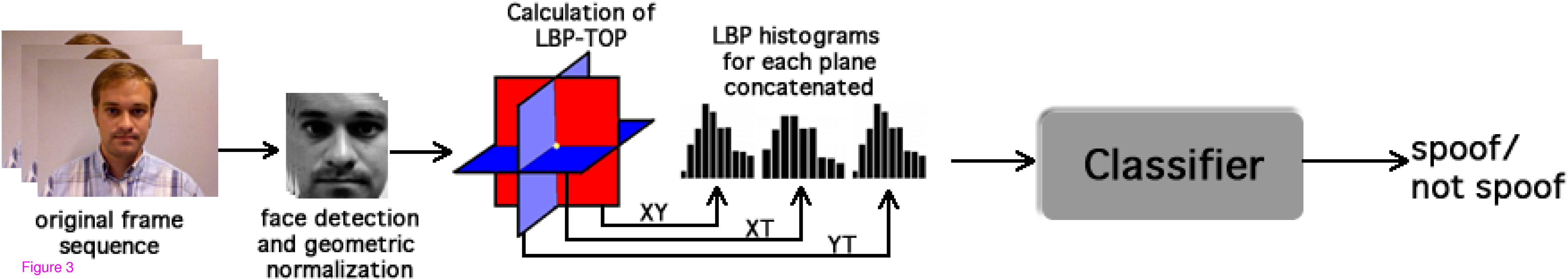

Calculation of LBP-TOP
LBP histograms for each plane concatenated
original frame sequence
face detection and geometric normalization
XY
XT
YT
Classifier
spoof/ not spoof

Figure 3

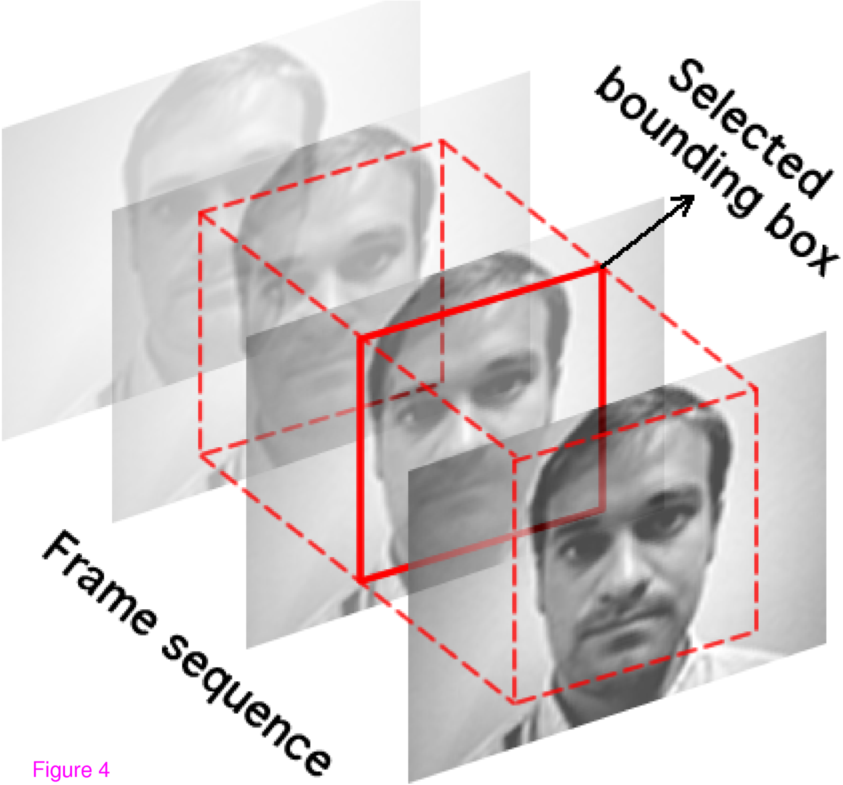


Figure 4

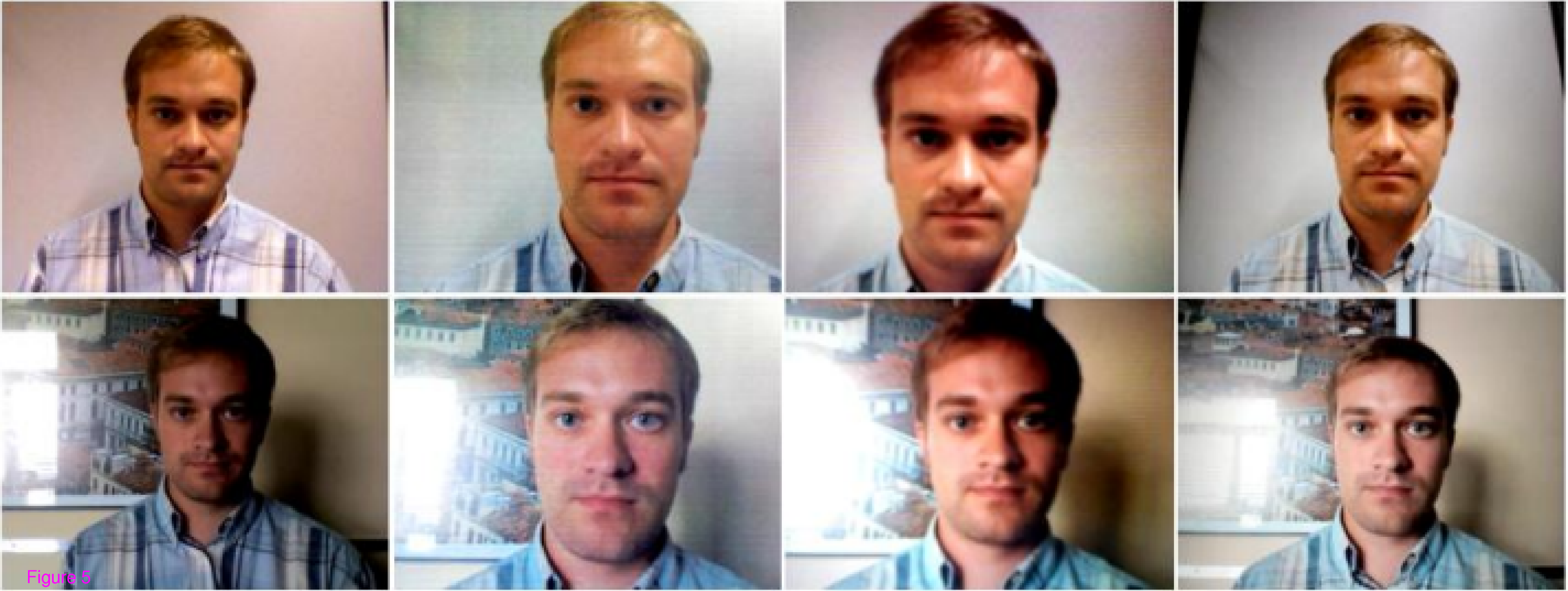


Figure 5

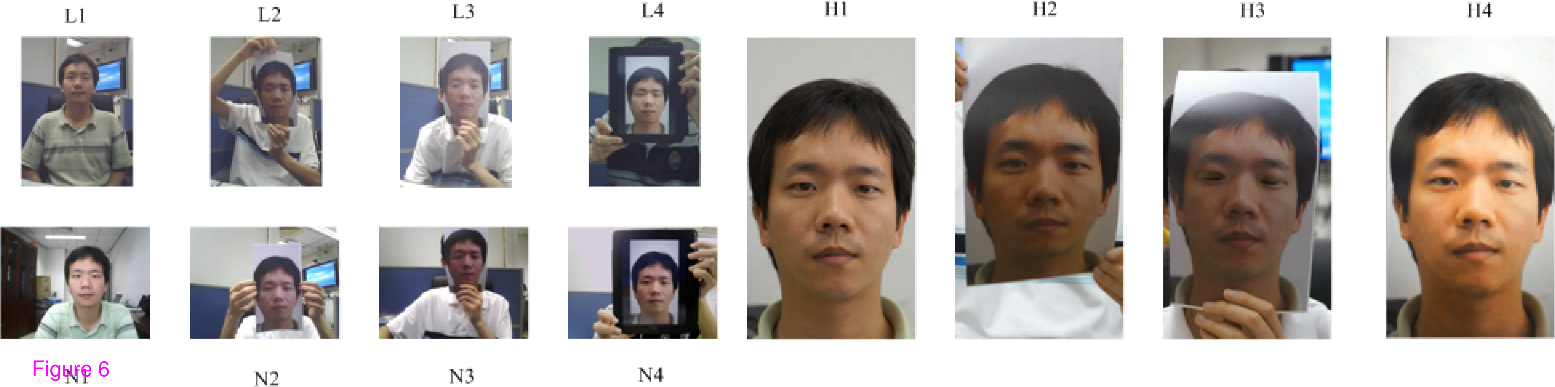


Figure 6

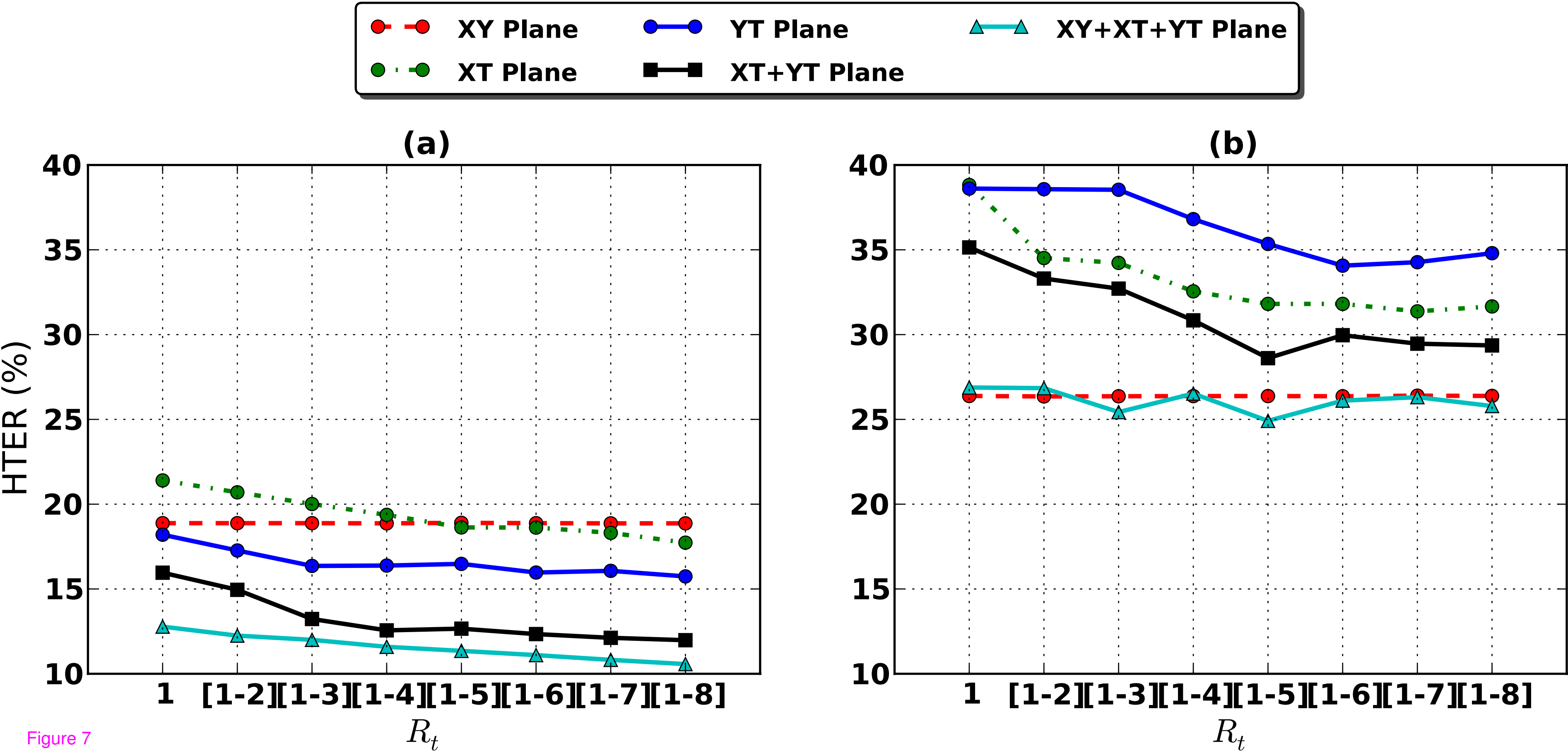

XY Plane
YT Plane
XY+XT+YT Plane
XT Plane
XT+YT Plane
(a)
(b)
HTER (%)
40
35
30
25
20
15
10
1
[1-2]
[1-3]
[1-4]
[1-5]
[1-6]
[1-7]
[1-8]
R_t

Figure 7

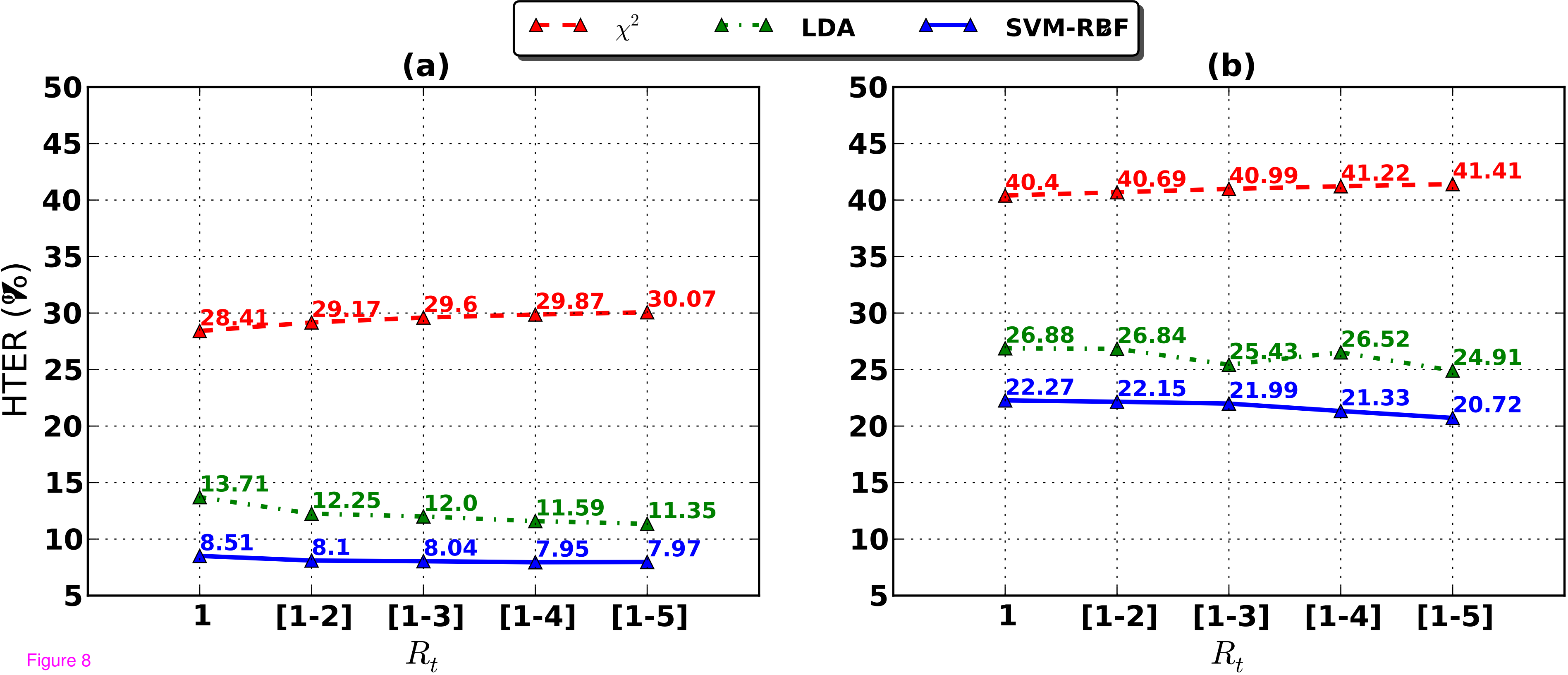

χ^2
LDA
SVM-RBF
(a)
(b)
HTER (%)
50
45
40
35
30
25
20
15
10
5
1
[1-2]
[1-3]
[1-4]
[1-5]
R_t
28.41
29.17
29.6
29.87
30.07
13.71
12.25
12.0
11.59
11.35
8.51
8.1
8.04
7.95
7.97
40.4
40.69
40.99
41.22
41.41
26.88
26.84
25.43
26.52
24.91
22.27
22.15
21.99
21.33
20.72

Figure 8

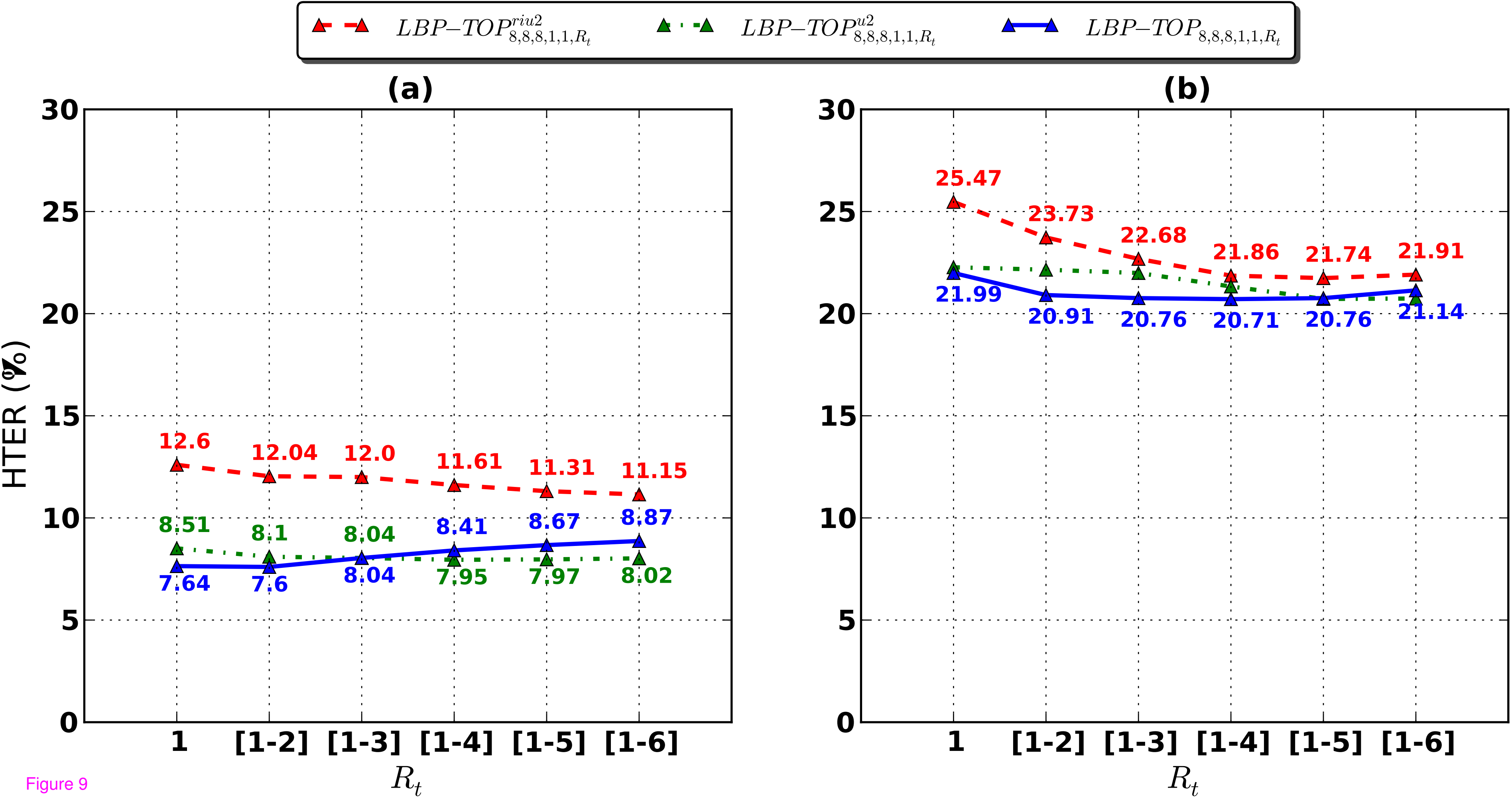

$LBP\text{–}TOP^{riu2}_{8,8,8,1,1,R_t}$
$LBP\text{–}TOP^{u2}_{8,8,8,1,1,R_t}$
$LBP\text{–}TOP_{8,8,8,1,1,R_t}$
(a)
(b)
HTER (%)
0
5
10
15
20
25
30
1
[1-2]
[1-3]
[1-4]
[1-5]
[1-6]
R_t
12.6
12.04
12.0
11.61
11.31
11.15
8.51
8.1
8.04
7.95
7.97
8.02
7.64
7.6
8.04
8.41
8.67
8.87
25.47
23.73
22.68
21.86
21.74
21.91
21.99
20.91
20.76
20.71
20.76
21.14

Figure 9

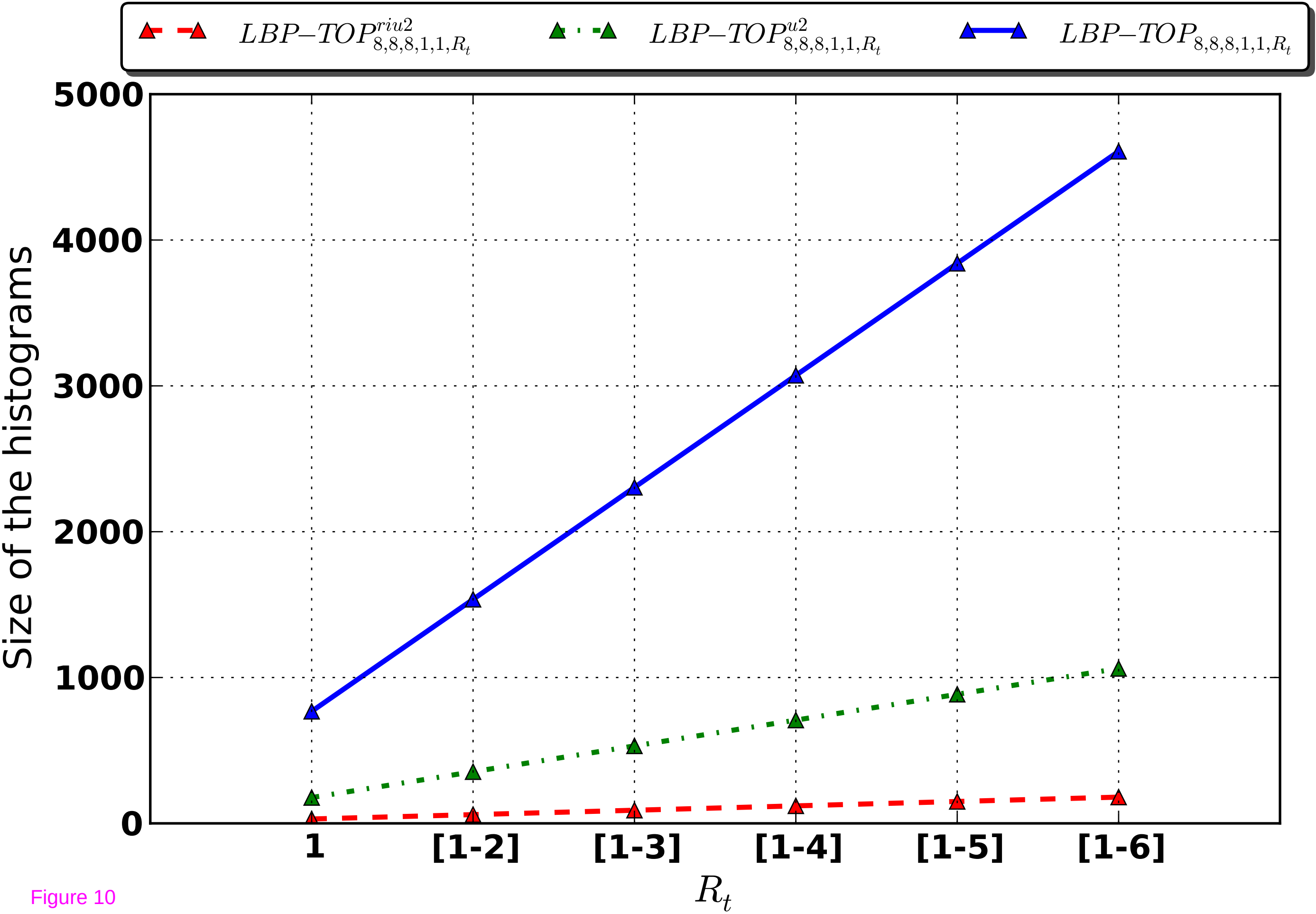

$LBP\text{-}TOP^{riu2}_{8,8,8,1,1,R_t}$
$LBP\text{-}TOP^{u2}_{8,8,8,1,1,R_t}$
$LBP\text{-}TOP_{8,8,8,1,1,R_t}$
Size of the histograms
5000
4000
3000
2000
1000
0
1
[1-2]
[1-3]
[1-4]
[1-5]
[1-6]
R_t

Figure 10

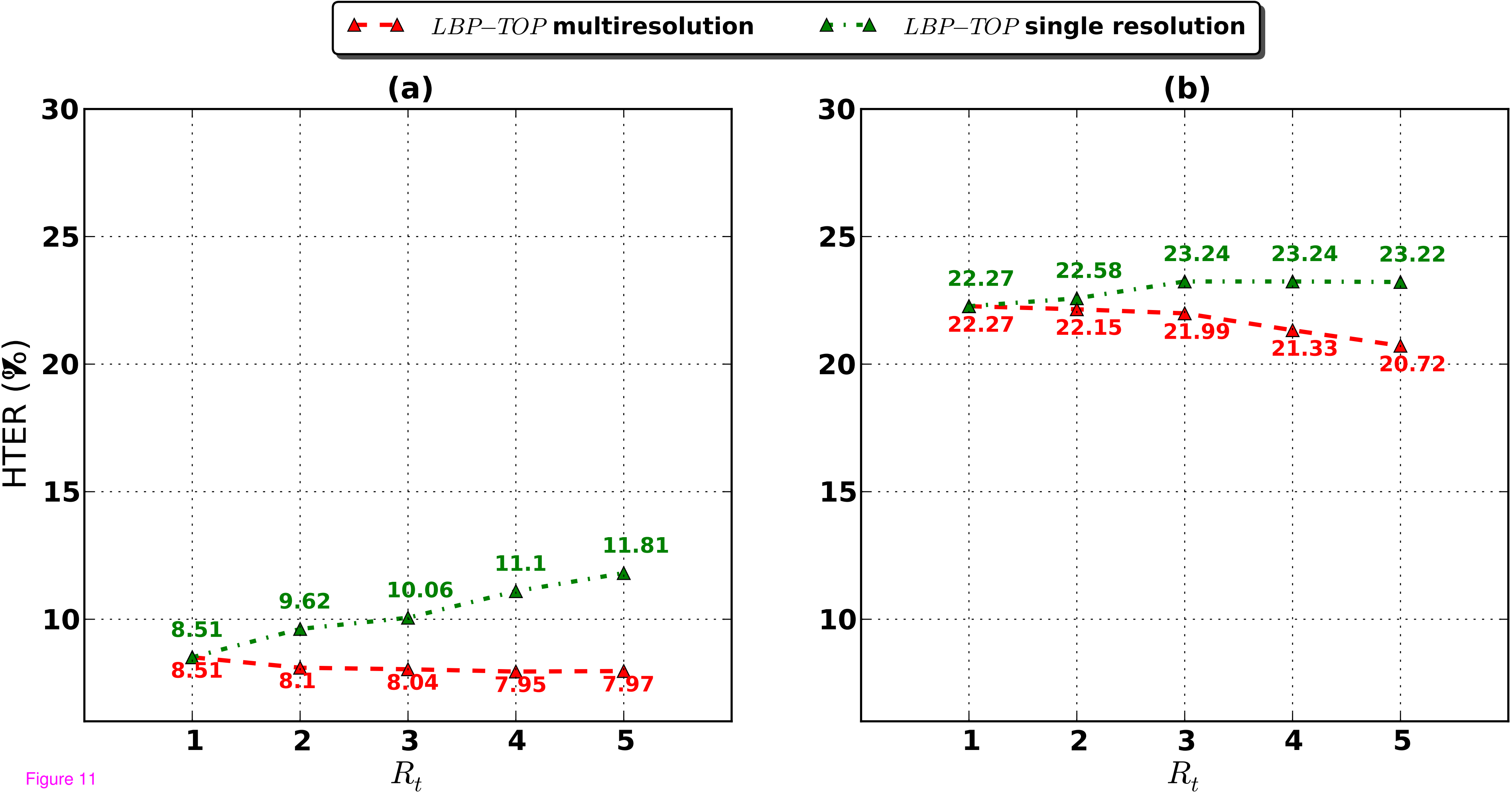
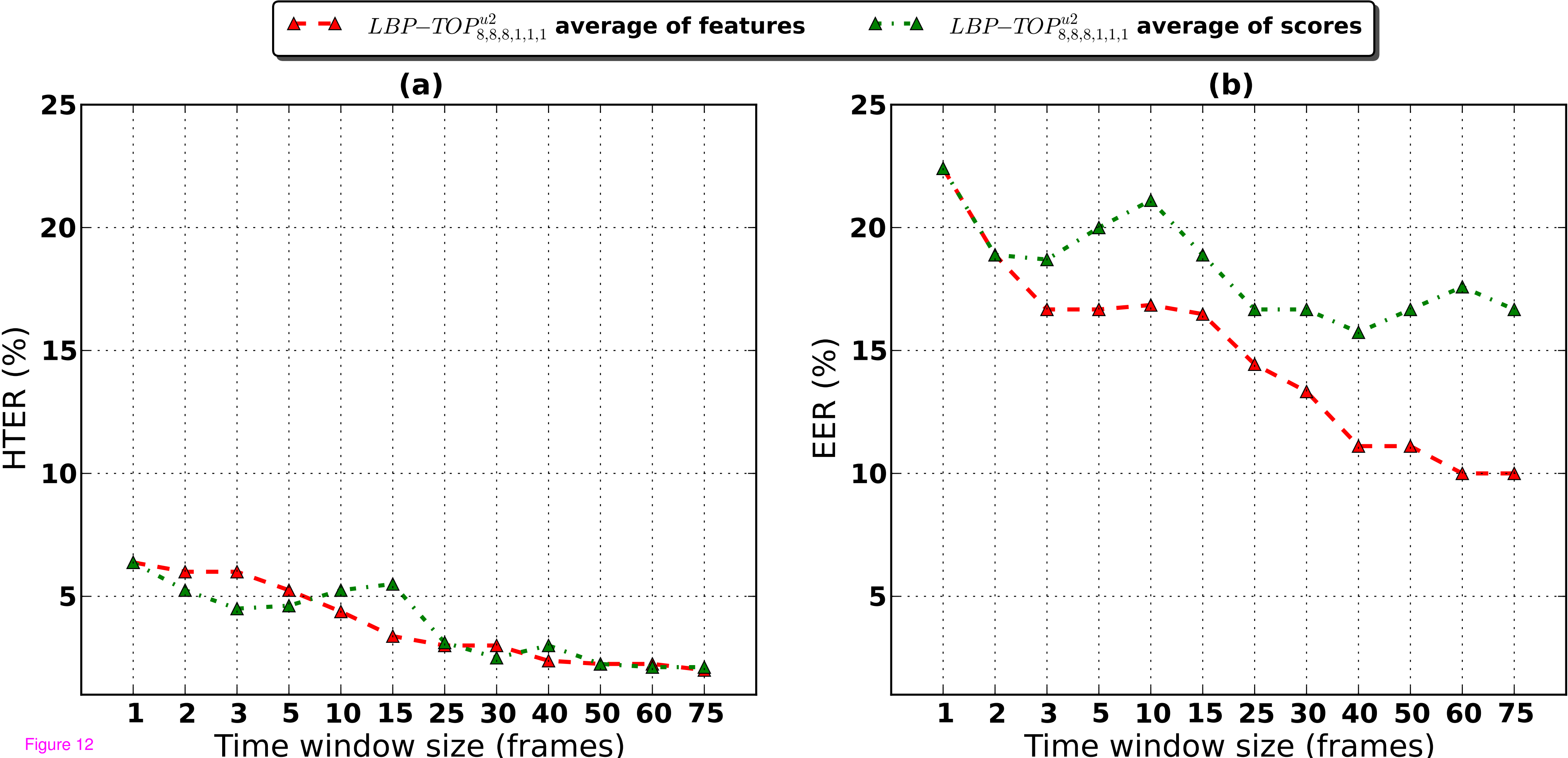

$LBP\!-\!TOP$ multiresolution
$LBP\!-\!TOP$ single resolution
(a)
(b)
HTER (%)
30
25
20
15
10
1
2
3
4
5
R_t
8.51
9.62
10.06
11.1
11.81
8.51
8.1
8.04
7.95
7.97
22.27
22.58
23.24
23.24
23.22
22.27
22.15
21.99
21.33
20.72

Figure 11



$LBP\text{–}TOP^{u2}_{8,8,8,1,1,1}$ average of features
$LBP\text{–}TOP^{u2}_{8,8,8,1,1,1}$ average of scores
(a)
(b)
HTER (%)
EER (%)
25
20
15
10
5
1
2
3
5
10
15
25
30
40
50
60
75
Time window size (frames)
Time window size (frames)

Figure 12

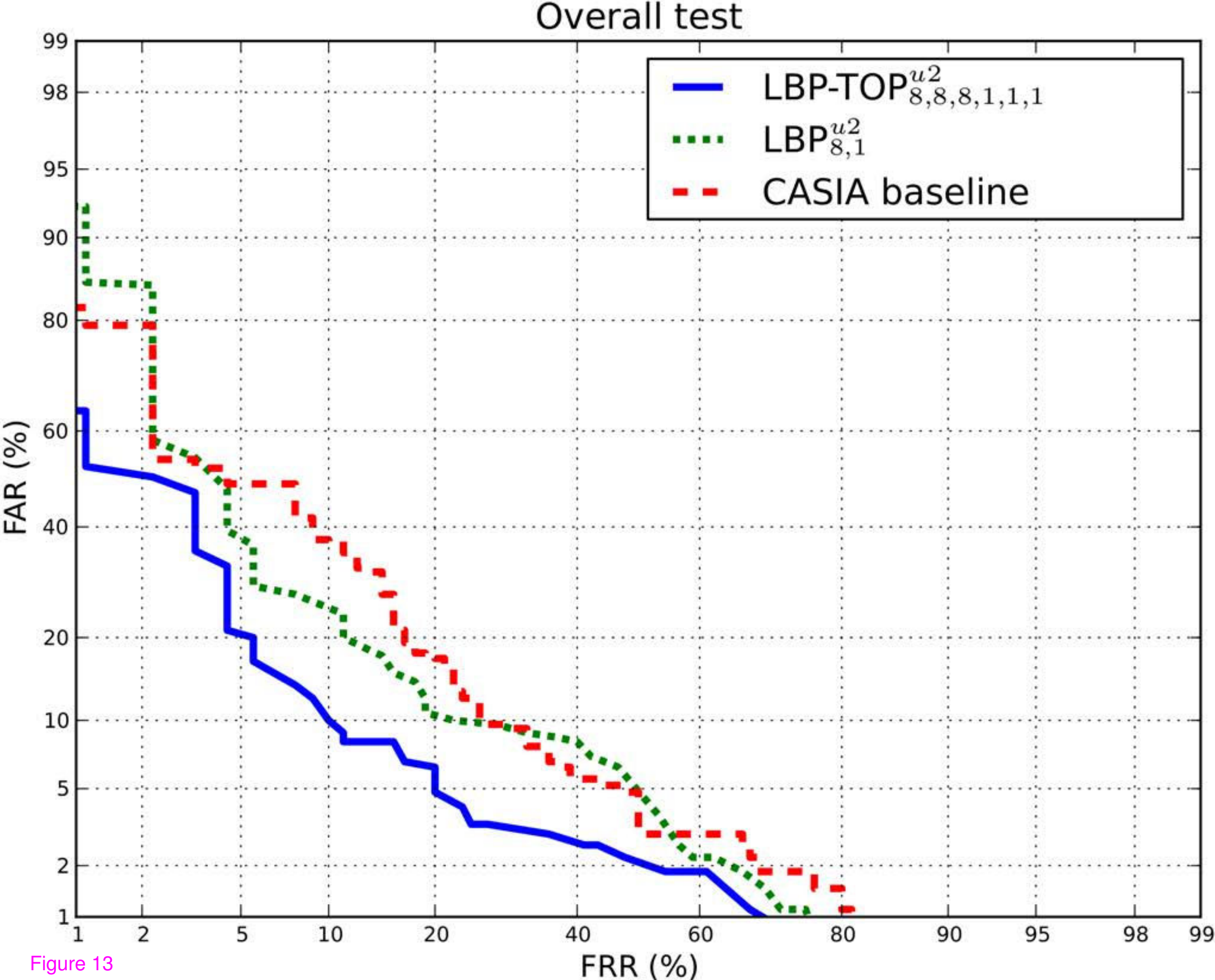

Overall test
LBP-TOP$^{u2}_{8,8,8,1,1,1}$
LBP$^{u2}_{8,1}$
CASIA baseline
FAR (%)
99
98
95
90
80
60
40
20
10
5
2
1
FRR (%)
1
2
5
10
20
40
60
80
90
95
98
99

Figure 13

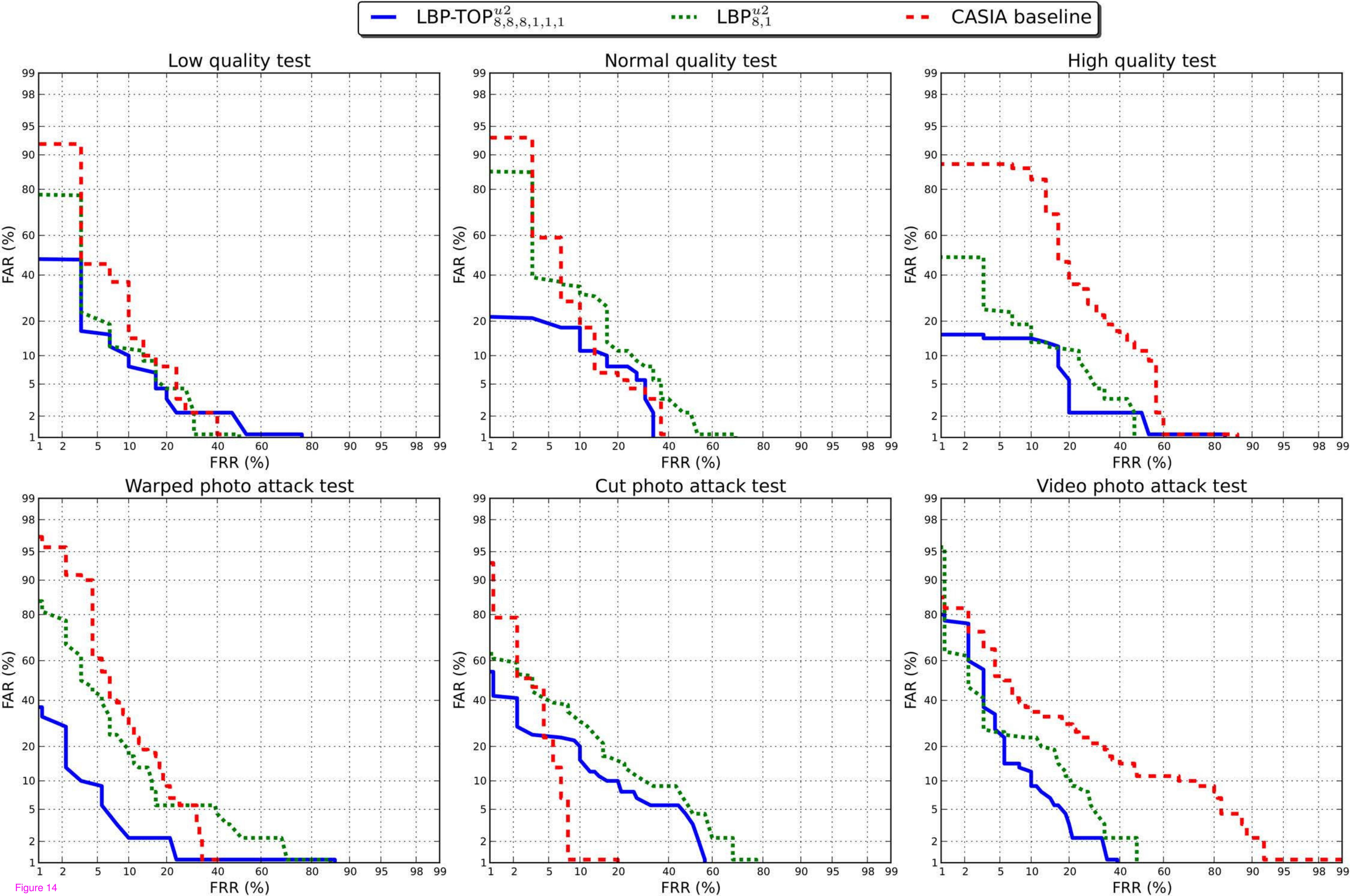

$\text{LBP-TOP}^{u2}_{8,8,8,1,1,1}$
$\text{LBP}^{u2}_{8,1}$
CASIA baseline
Low quality test
Normal quality test
High quality test
Warped photo attack test
Cut photo attack test
Video photo attack test
FAR (%)
FRR (%)
1
2
5
10
20
40
60
80
90
95
98
99

Figure 14