

# NeuroBiometric: An Eye Blink Based Biometric Authentication System Using an Event-Based Neuromorphic Vision Sensor

Guang Chen, *Member, IEEE*, Fa Wang, Xiaoding Yuan, Zhijun Li, *Senior Member, IEEE*,  
Zichen Liang, and Alois Knoll, *Senior Member, IEEE*

**Abstract**—The rise of the Internet and identity authentication systems has brought convenience to people’s lives but has also introduced the potential risk of privacy leaks. Existing biometric authentication systems based on explicit and static features bear the risk of being attacked by mimicked data. This work proposes a highly efficient biometric authentication system based on transient eye blink signals that are precisely captured by a neuromorphic vision sensor with microsecond-level temporal resolution. The neuromorphic vision sensor only transmits the local pixel-level changes induced by the eye blinks when they occur, which leads to advantageous characteristics such as an ultra-low latency response. We first propose a set of effective biometric features describing the motion, speed, energy and frequency signal of eye blinks based on the microsecond temporal resolution of event densities. We then train the ensemble model and non-ensemble model with our NeuroBiometric dataset for biometrics authentication. The experiments show that our system is able to identify and verify the subjects with the ensemble model at an accuracy of 0.948 and with the non-ensemble model at an accuracy of 0.925. The low false positive rates (about 0.002) and the highly dynamic features are not only hard to reproduce but also avoid recording visible characteristics of a user’s appearance. The proposed system sheds light on a new path towards safer authentication using neuromorphic vision sensors.

**Index Terms**—Biometrics, biometric authentication, event-based vision, neuromorphic vision.

## I. INTRODUCTION

Manuscript received May 10, 2020; revised July 4, 2020; accepted July 22, 2020. This work was supported by the National Natural Science Foundation of China (61906138), the National Science and Technology Major Project of the Ministry of Science and Technology of China (2018AAA0102900), the Shanghai Automotive Industry Sci-Tech Development Program (1838), the European Union’s Horizon 2020 Research and Innovation Program (785907), and the Shanghai AI Innovation Development Program 2018. Recommended by Associate Editor Chenglin Liu. (*Corresponding author: Zhijun Li*.)

Citation: G. Chen, F. Wang, X. D. Yuan, Z. J. Li, Z. C. Liang, and A. Knoll, “NeuroBiometric: An eye blink based biometric authentication system using an event-based neuromorphic vision sensor,” *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 1, pp. 206–218, Jan. 2021.

G. Chen is with the Tongji University, Shanghai 200092, China, and also with the Technical University of Munich, Munich 80333, Germany (e-mail: guang@in.tum.de).

F. Wang, X. D. Yuan, and Z. C. Liang are with the Tongji University, Shanghai 200092, China (e-mail: 18007873581@163.com; xiaodingyuan.tj@gmail.com; m15801781928@163.com).

Z. J. Li is with the University of Science and Technology of China, Hefei 230026, China (e-mail: zjli@ieec.org).

A. Knoll is with the Technical University of Munich, Munich 80333, Germany (e-mail: knoll@in.tum.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JAS.2020.1003483

BIOMETRIC authentication has long been studied by analyzing people’s biological and behavioral characteristics. Distinct biological characteristics from various sensors have been used as a part of effective biometrics in practical application scenarios, such as in privacy preservation and identity recognition. Present systems based on fingerprint scanners [1] or cameras [2] have enjoyed great popularity, bringing rapid development in related researches. But some risks still exist, including the risk that biometrics based on static physical appearance are easier to fake compared with those with highly dynamic features.

Inspired by electroencephalogram-based robotics control [3]–[5], some research [6], [7] has focused on the subtle physiological behavior caused by bio-electric currents, especially with eye movement. Clearly, eye movement features consist of implicit and dynamic patterns, and are harder to artificially generate. Thus, the system can be less likely to be spoofed. These implicit dynamic patterns of eye blink make them unique in identity recognition, and are difficult to be captured at a distance. In addition, eye blink features also have several other applications, like in eye detection and tracking [7] or face detection [8]. Another crucial factor is the sensor. At present, an iris scanner [9] could be the safest but its expensive cost prevents it from wider application. Another method studied by [10] aims to utilize the electro-oculogram (EOG) with the Neurosky Mindwave headset to study the detected voltage changes caused by blinks. However, wearing the devices poses a certain interference with spontaneous blinking, which hinders user convenience and exacerbates system complexity.

In this paper, we introduce a neuromorphic vision-based biometric authentication system, incorporating microsecond level event-driven eye-blinking characteristics. The main device in our system is a DAVIS sensor [11], a bio-inspired dynamic vision sensor that works completely differently with conventional ones. While a conventional camera records the entire image frames at a fixed frame rate, a DAVIS sensor only captures and transmits the intensity changes of pixels asynchronously caused by any subtle motions in a scene. Once a change reaches a certain threshold, the pixel will instantly emit a signal called an event. Compared with other biometric authentication sensors, it has no direct contact with the subjects and is purely passive that it only captures triggered events instead of radiating rays. In this work, the processing

procedures are as follows: we first sum up the events for each recording with a sliding window and then fit a median curve to describe dynamic features in time and frequency domain. After that, we conduct a feature selection and train an ensemble model and a non-ensemble model for authentication. And the dataset<sup>1</sup> will be available online to encourage the comparison of any authentication method with this work.

The major contributions of our work can be summarized in the following three aspects:

1) Proposing the first-ever neuromorphic sensor-based biometric authentication system. It is capable of capturing subtle changes of human eye blinks in microsecond level latency which traditional CMOS cameras can not.

2) Constructing a series of new features out of our biometric data (eye blinks): duration features, speed features, energy features, ratio features, and frequency features.

3) The system provides effective feature selection and identity recognition approaches in authentication, and it is proved to be robust and secure with a very low false-positive rate.

## II. RELATED WORK

Existing explicit and static biometric authentication features could be less difficult to mimic and may pose a threat to people's privacy and security. To pursue more secure systems, researchers proposed several implicit biometrics identification approaches [12], [13]. Among all the bio-signals, eye movement signals are relatively simpler to process in terms of data acquisition, and usually utilizes EOG and video-oculography (VOG) to study biometric authentication systems.

### A. EOG Based Methods

Researches involving EOG based methods mainly transfer electrode change information from eyeball rotation into biometric features with a head-mounted detector. The features are further utilized to train learning models for the biometric authentication system. Reference [10] states that the biometric features that describe eye blink energy are important for the identification of subjects. In [14], they adopt Shannon entropy (SE) to extract features for fast EEG-based person identification. [15] proposes a discrete wavelet transform-based feature extraction scheme for the classification of EEG signals. Reference [16] uses the Hilbert transform to extract the average phase-locking values, the average instantaneous amplitudes and spectral power from EEG signals. In recent works, authentication of [17] is based on features extracted from eye blinking wave forms. Features used by [6] are made up of event-related potentials generated from rapid serial visual presentation diagrams and morphological features extracted from eye blinking data. In addition to the feature extraction process, feature selection is also crucial to ensure the accuracy and effectiveness of the system. Reference [18] applies feature selection approaches like discrete Fourier transform coefficients in authentication scenarios. The authentication system based on EOG often has better

accuracy. However, it is only acceptable in very limited scenarios like in the hospital because all subjects have to wear a head-mounted device for long-duration records, which is not practical in daily use.

### B. VOG Based Methods

VOG based methods mainly achieve identity authentication via eye detection and iris recognition with images clipped from videos. In the work [19], the system detects iris centers on low-resolution images in the visible spectrum with an efficient two-stage algorithm. Reference [20] develops a general framework based on fixation density maps (FDMs) for the representation of eye movements recorded under the influence of a dynamic visual stimulus with commercial iris recognition devices. Another work [21] proposes and trains a mixed convolutional and residual network for the eye recognition task, which improves the system's anti-fake capability. Reference [22] provides a non-contact technique based on high-speed recording of the light reflected by eyelid in the blinking process for feature extraction and authentication. Reference [23] evaluates features generated from eye fixations and saccades with linear discriminant analysis (LDA). But identification and verification based on VOG have an unavoidable defect. Common VOG sensors work at a fixed frame rate so it is difficult to achieve a very high time resolution. This can result in motion blur when capturing highly dynamic motions like a blink. Also, capturing and transmitting the entire image frame continuously lowers the efficiency that the system can achieve.

### C. Neuromorphic Vision Methods and Applications

The DAVIS sensor is an emerging bio-inspired neuromorphic vision sensor that records the intensity changes of pixels caused by motions in a scene asynchronously. The differences between conventional cameras and neuromorphic vision sensors are depicted in Fig. 1. Compared with a conventional frame-based camera, it has significantly higher time resolution and lower latency [24]–[26] places a DAVIS sensor on a moving pan-tilt unit on a robotic platform to

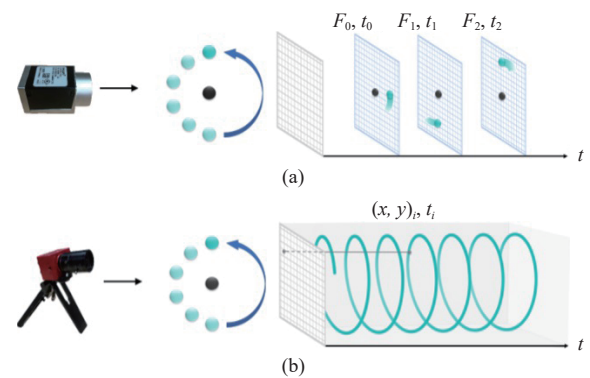


Fig. 1. The differences between conventional cameras and neuromorphic vision sensors. A green ball is rotating around a centering black ball. (a) A conventional camera captures all pixel intensities at a fixed frame rate, e.g., black ball and green ball with motion blur. (b) A neuromorphic vision sensor captures only intensity changes caused by the motions of the green ball asynchronously.

<sup>1</sup> <https://github.com/ispc-lab/NeuroBiometrics>

recognize fan-like objects. Reference [27] proposes a method to detect faces based on eye blink signals recorded with a DAVIS sensor. A comprehensive review of applications with the DAVIS sensor is addressed in [28], [29].

### III. APPROACH

In this section, we first illustrate the neuromorphic biometrics dataset. Then we introduce the signal processing techniques and the extraction process of biometric features.

#### A. Neuromorphic Biometrics Dataset

To study the biometric authentication system with an event-based neuromorphic vision sensor, we collected a new dataset with eye blink signals recorded by the DAVIS sensor named NeuroBiometric dataset. The DAVIS sensor we used is DAVIS346, which has a resolution of  $346 \times 260$  pixels, a temporal resolution of  $1 \mu\text{s}$  and an outstanding dynamic range (up to 140 dB). We have 45 volunteers (of whom, 23 are men and 22 are women) to participate in our recording. All the volunteers are in a normal psychological and physiological state.

The dataset collection is conducted during the day in an indoor environment with natural light exposure. Here, the light intensity is approximately 1200 LUX. The value is measured by a photometer and we observe that the light intensity has limited influence on the authentication quality as long as the dataset is collected under the natural light of the day. This is because proper event sequences can be generated by adjusting the aperture through the camera or the contrast sensitivity of the event camera (an adjustable parameter relative to the event threshold) through the software kit, even though the number of triggered events has a positive correlation with light intensity. Not to mention that the event camera has a higher dynamic range compared to traditional camera (more than 120 dB versus 60 dB) and the dynamic range of the natural light intensity at the same place is not very high during a day.

Subjects are instructed to sit in front of a DAVIS sensor with their head on a bracket for stabilization, then gaze ahead and blink spontaneously. Only the facial region that includes the eyebrow and eye is recorded by the DAVIS sensor. An additional 60 seconds of break time after every 120 seconds recording is implemented to avoid unusual eye blinks due to potential fatigue and distraction. Each subject has performed at least 4 blinking sessions. In this work, we select 480 seconds (8 minutes) of valid spontaneous eye blink data for each subject. These blinks rate varies from 10 to 36 times every minute, and the duration ranges from 250 to 500 ms.

#### B. Neuromorphic Signal Processing

1) *Presentation of Event and Event Density*: The DAVIS sensor is a high-speed dynamic vision sensor that is sensitive to intensity changes. Formally, each pixel circuit of a DAVIS sensor tracks the temporal contrast defined as  $c = d(\log I)/dt$ , which is the gradient of the log-intensity of the incoming light. Increase of light intensity generates an ON event (or positive event), an OFF event (or negative event) is caused by reduced light intensity. An event will be triggered when the

temporal contrast exceeds a threshold in either direction. Each event is represented by a 4-dimensional tuple  $(ts, x, y, p)$ , where  $ts$  is the timestamp when an event is triggered at pixel coordinates  $(x, y)$ ;  $p$  is the polarity of the event (i.e., brightness increase (“ON”) or decrease (“OFF”). The DAVIS sensor has a microsecond level accuracy, and a typical eye blink can generate 50 K–200 K events.

The original event data stream is a sequence of events sorted by time, in which every event is represented by a 4-d tuple mentioned above. The sequence of events could be described as a series of events  $e_1, e_2, \dots, e_n$  with corresponding timestamps  $t_1, t_2, \dots, t_n$ . We slide a particular window along the time axis to select a certain number of events and define event density as the event counts in the window. The timestamp is marked by the absolute timestamp of the first event included. Our experiment is conducted based on the curves between event density and the timestamp. Set the event density of timestamp  $T$  as  $D_T$ , and the windows size as  $W$ ,  $D = \sum_{x=i}^j e_x, T \leq t_i < t_{i+1} < \dots < t_j < T + W$ . Due to the high precision of the DAVIS sensor, we calculate the event density in a window size of 10 ms. Figs. 2 and 3 show an example of eye blink event density stream.

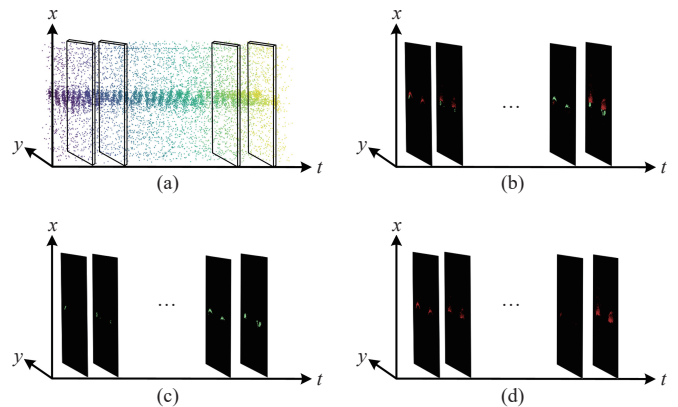


Fig. 2. (a) A 60-second raw event stream of eye blinks. (b) Event frames sliced from raw blink stream, where positive events are in red and negative events in green. (c) ON (positive) event frames sliced from raw blink stream. (d) OFF (negative) event frames sliced from raw blink stream (better viewed in color).

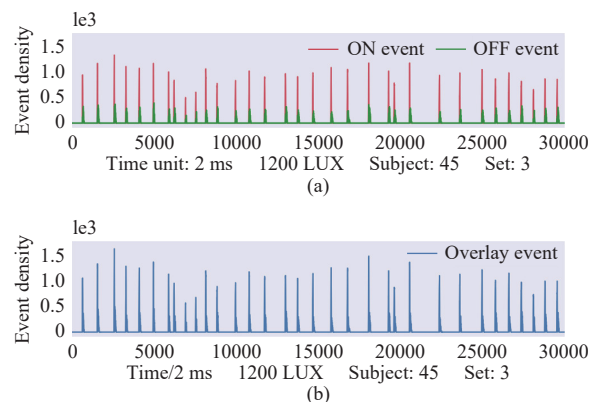


Fig. 3. Event density of a series of eye blinks, in the type of OFF event (green line), ON event (red line), and overlay event of both (blue line). Those curves are denoised (better viewed in color).

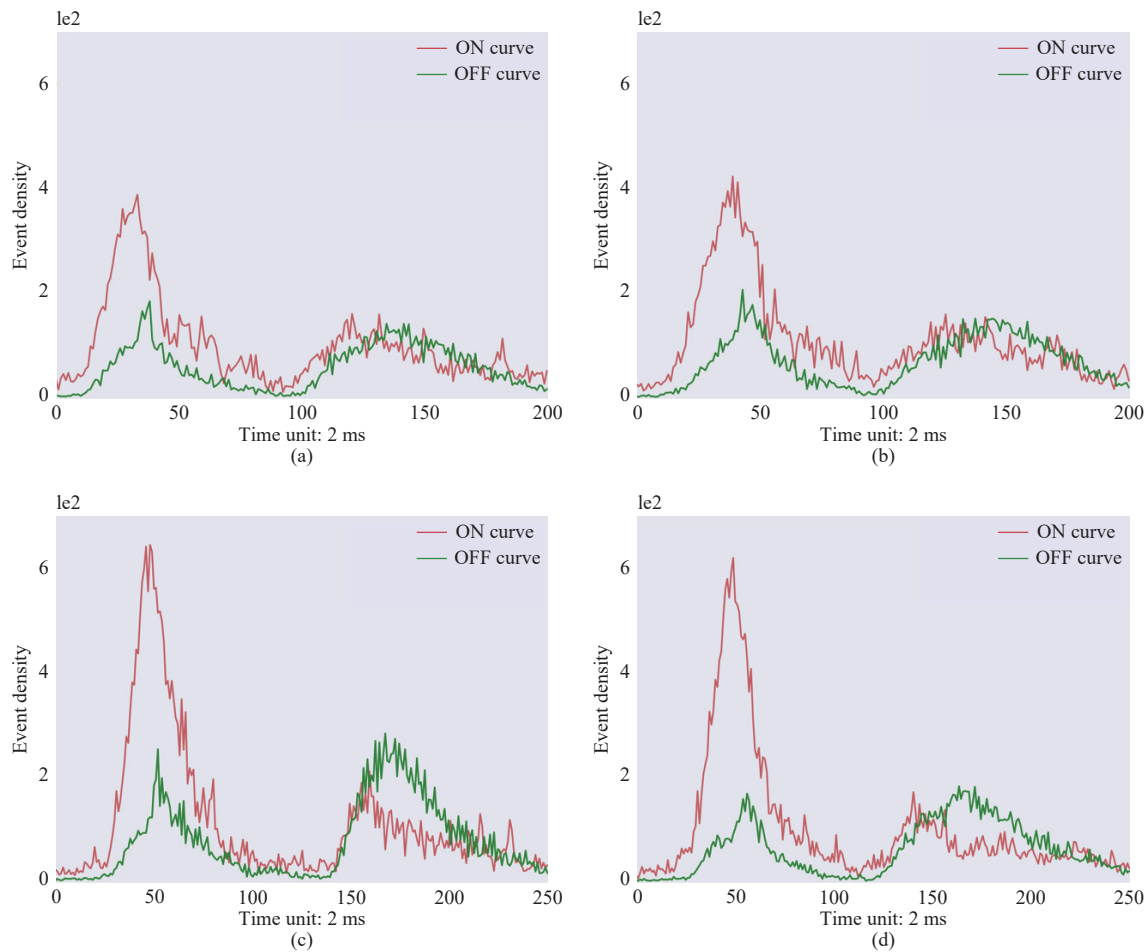


Fig. 4. Original eye blink signals of (a) subject A recorded early; (b) subject A recorded late; (c) subject B recorded early; (d) subject B recorded late (better viewed in color).

2) *Filtration and Curve Smoothing*: Fig. 4 shows the original blink signal of different subjects recording at different times. We observe that features like maximum peak value and blink duration and so on are different among subjects. The maximum peak value and the blink duration of the curves in both Figs. 4(c) and 4(d) are larger than those in 4(a) and 4(b). However, the curve is not smooth enough to extract features so that we propose a signal processing algorithm named filtration and curve smoothing to smooth the curve. Fig. 5(a) shows a sequence of event density data with an eye blink. Ideally, few events are supposed to be triggered when there is no eye blink. However high frequency and low energy light fluctuation cannot be averted completely. In the course of our experiment, noises are even stronger under an incandescent light condition. We apply neighbor pixel communication (NPC) filtering as a filtration algorithm to the DAVIS sensor. The NPC structure is designed to judge the validity of pixels activity through the communication between its adjacent pixels. The communication of adjacent pixels can be described as a  $N \times N$  array, where  $N$  is odd. The validity of the central pixel is determined if the total amount of events in the  $N \times N$  array exceeds a defined threshold of  $t$  (in the order of pixel). In this work,  $N = 5$  and  $t = 8$ .

Fig. 5(b) shows the result of the NPC filtration algorithm as a comparison to the initial event density stream as Fig. 5(a).

Additionally, a simple smoothing algorithm is also introduced here, which replaces each point of the event density curve with the average of  $w$  nearby points, where  $w$  is a positive integer called the smooth width. In our work,  $w = 5$ , and the time unit is 2 ms. Fig. 5(c) shows the effect of our curve smoothing method. In combination of both the event NPC filtration and the smoothing method, Fig. 5(d) shows the event density curve for the biometric feature extraction.

### C. Neuro-Biometric Feature Extraction

Though eye blink frequency and speed can be significantly influenced by emotion, fatigue, level of focus, age and light condition, spontaneous eye blink under a normal physical and emotional state and moderate illumination intensity can be relatively stable to one specific individual, especially when focusing on the pattern of eyelid movement during a single eye blink.

Fig. 3 shows an event density stream of 60 s of eye blink data. Since the DAVIS sensor can only detect changes of light intensity, few events would be triggered when eyes are gazing. When an eye blink starts, the upper eyelid moving downwards and activates most pixels around the eyelid and the eyelash. In this work, eye blink dynamic features extracted from eye blink signals are generally based on the active event density during eye blinks. Just as Fig. 6 shows, 4 curves

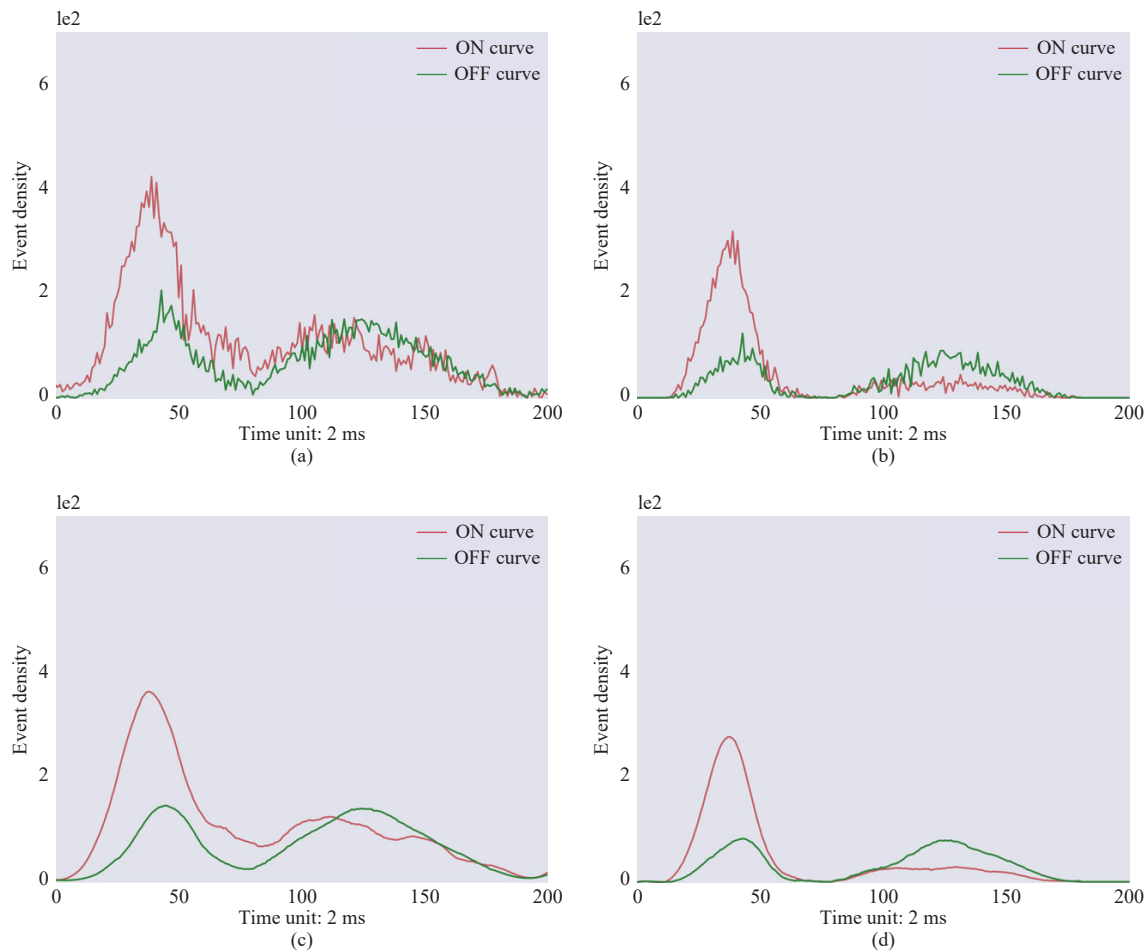


Fig. 5. Eye blink event density curves with (a) neither filtration nor smoothing; (b) NPC filtration; (c) a smoothing method with a window size of 10 ms; (d) both filtration and smoothing method (better viewed in color).

(named ON, OFF, Overlap, and Splice curve, respectively) between event density and timestamp are generated to describe the active event density. The ON curve represents its event density and is the summation of ON events in the sliding window while the OFF curve sums only OFF events as its event density. Overlap events equals the summation of the ON curve and OFF curve. The splice curve is a joint curve between the ON curve during the closure and the OFF curve during the re-opening.

Regarding the single eye blink event density stream in Fig. 7(a), the curve reaches two peaks: one during eye closing and one during eye re-opening. The lowest point between two peaks implies a fleeting state of closed eye between eye closing and eye re-opening. The biological principle lying behind the two peaks is brightness difference between eyelid, iris and sclera. The sclera is brighter than the upper eyelid but the upper eyelid is much brighter than the iris. Both ON and OFF events are triggered during the whole eye-blinking movement. The iris can be regarded as a black disk. The brightness change caused by the upper eyelid transient movement can be regarded as a horizontal line, the length of which (iris plus sclera) increases until the middle after which it reduces. This leads to two peaks in the blinking pattern. And the first peak is higher than the second one due to the faster rate of eye closure.

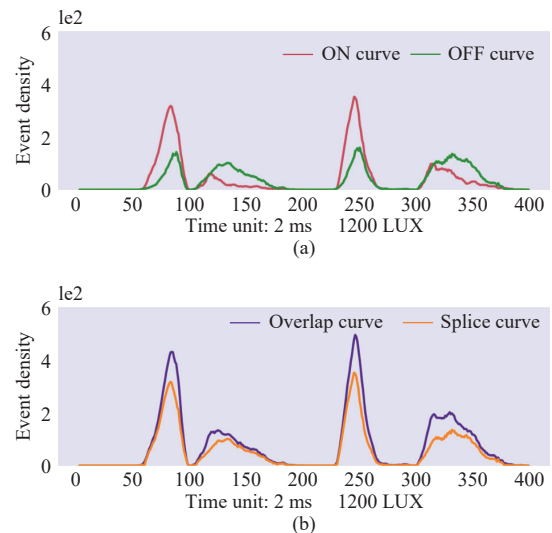


Fig. 6. (a) ON curve: ON event density stream of eye blink. OFF curve: OFF event density stream of eye blink. (b) Overlap curve: overlap stream of eye blink contains both OFF and ON events. Splice curve: Spliced stream of eye blink includes only OFF events during eye closing and ON events during eye re-opening (better viewed in color).

1) *Eye Blink Stage Detection*: Eye blink stage can be intuitively divided into eye closing and eye re-opening. These

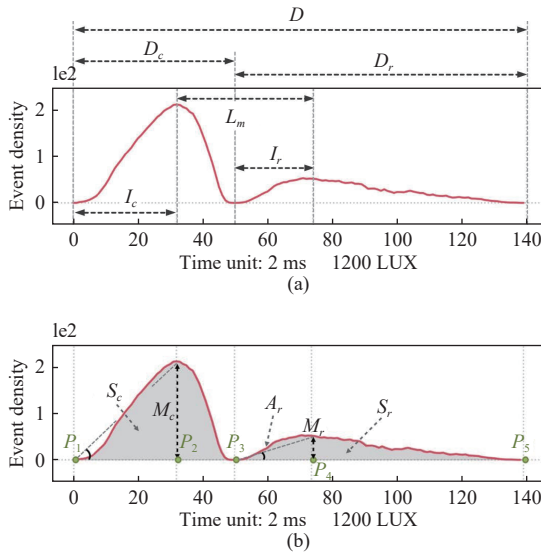


Fig. 7. The visualization of feature symbols illustration in a typical eye blink event density stream (better viewed in color).

two stages of eye movement have many similar and complementary features. The physical motion of the eye blink can be divided into eye closing and eye reopening. During eye closing, the upper eyelid moves towards the lower eyelid. Negative events that are caused by the eyelid moving downwards generally appear in higher latitudes than positive events do. Then after a short period of the eye closed state, the eyelid moves upwards again and causes more positive events. Based on the opposite movement of the eyelid, eye closing, and eye reopening could be distinguished by the distribution of event polarity triggered by the DAVIS sensor.

With eye closing and eye re-opening, one thing in common is that the speed of eyelid movement raises from 0 to a maximum, and then slows down to 0 again. Thus, there must be a maximum peak of event density during both the eye closing stream and eye re-opening stream. Significantly, both of the maximum peaks can be influenced by the width and thickness of the eyelid or the length of an eyelash.

In addition to what is mentioned above, we must set a proper boundary of the eye blink event density curve. We propose that the threshold for a valid eye blink curve contains 3% – 100% intensity of the maximum value in a series of eye blink data. Event density under the threshold would be considered as invalid noise.

In this work, the five points mentioned in Fig. 7(b), namely, the start of eye closing ( $P_1$ ), maximum peak of the eye closing ( $P_2$ ), middle of eye closed state ( $P_3$ ), maximum peak of eye re-opening ( $P_4$ ), and end of eye re-opening ( $P_5$ ), are critical to the feature extraction of the eye blink event density stream.

2) *Eye Blink Features*: A set of 21 basic features extracted from the eye blink event density stream (henceforth referred to as *eye blink stream*) are used for further derivation and feature selection. Basic features are grouped into 5 groups. A detailed description of basic features can be found in Table I. The explanations of these features are as below:

a) *Duration features*: Duration features describing the time scales of different eye blink stages are shown in Fig. 7(a). The

TABLE I  
SUMMARY OF DYNAMIC FEATURES EXTRACTED FROM EYE BLINK EVENT DENSITY STREAM

Group	Symbol	Description
Duration	$D$	Duration of eye blink
	$D_c$	Closure duration of eye blink
	$D_r$	Reopening duration of eye blink
	$L_m$	Projection on x-axis between the eye closing peak and eye reopening peak
	$I_c$	Projection on x-axis between the start of eye closing and eye closing peak
Energy	$I_r$	Projection on x-axis between the start of eye reopening and eye reopening peak
	$M_c$	Amplitude of eye closing peak
	$M_r$	Amplitude of eye reopening peak
	$A_c$	Angle between eye closing peak and x-axis
	$A_r$	Angle between eye reopening peak and x-axis
Ratio	$S_c$	Area between eye closing curve and x-axis
	$S_r$	Area between eye reopening curve and x-axis
	$R_L$	The ratio of $L_m$ to $D$
Frequency	$R_M$	The ratio of $M_c$ to $M_r$
	$R_S$	The ratio of $S_c$ to $S_r$
	$F_p$	FFT peak amplitude
	$HF_p$	FFT half-waveform peak amplitude
	$F_a$	FFT arithmetic mean amplitude
Speed	$F_{min}$	FFT waveform minimum
	$R_f$	Ratio of $F_p$ to $HF_p$
	$\Sigma_f$	Sum of $F_p$ and $HF_p$
Speed	$T_m$	Mean of adjacent events time interval on pixel level
	$T_v$	Variance of adjacent events time interval on pixel level

duration of an eye blink is defined as  $D$ , where  $D_c$  and  $D_r$  are the duration of eye closing ( $P_1$  to  $P_3$ ) and eye re-opening ( $P_3$  to  $P_5$ ). In more detail,  $I_c$  is the projection on x-axis from start of eye closing ( $P_1$ ) to the maximum peak of eye closing ( $P_2$ ), while  $I_r$  is the projection on the x-axis from eye closed state ( $P_3$ ) to the maximum peak of eye re-opening ( $P_4$ ). Another duration feature calculates the projection on the x-axis between the eye closing peak and re-opening peak defined as  $L_m$ .

b) *Energy features*: Energy features of an eye blink stream contain amplitudes, area and tangents in the curve. We define  $M$  as amplitude, where the amplitude of the eye closing and eye re-opening maximum peak is defined as  $M_c$  and  $M_r$ . Tangent value  $A_c$ ,  $A_r$  is defined as  $M_c/I_c$ ,  $M_r/I_r$ .  $S_c$  calculates the area between the eye closing stream and x-axis, while  $S_r$  calculates the area between the eye re-opening stream and x-axis. Illustration can be found in Fig. 7(b).

c) *Ratio features*: Ratio features describe proportions of different types of features. The proportion of time  $R_L$  is defined as  $L_m/D$ , while proportion of amplitude  $R_m$  is defined as  $M_c/M_r$  and area ratio  $R_s$  is defined as  $S_c/S_r$ .

d) *Frequency features*: Frequency features are calculated from the frequency domain transformed from an eye blink stream using a fast Fourier transform (FFT) algorithm. Peak amplitude and half-waveform peak amplitude are defined as

$F_p$  and  $HF_p$ .  $F_a$  is the arithmetic mean amplitude.  $F_{\min}$  describes the waveform minimum. Ratio  $R_f$  is defined as  $F_p/HF_p$ , and the sum  $\sum_f$  calculates the sum of  $F_p$  and  $HF_p$ .

e) *Speed features*: The features listed above are generally based on the event density stream of eye blinks, while speed features are based on the raw data recorded by the DAVIS sensor. Regarding a pixel identified by a 2-dim coordinate  $(x,y)$ , events located on that pixel are likely to be triggered more than once in a period (like 1 ms). By assembling these events head-to-tail in sequence, the timestamp intervals of these events indicate how fast the lightening condition changes on a pixel-level. The parameters related to the distribution of time interval are regarded as speed features of the human eye blink. The mean and variance of adjacent events time intervals on pixel-level are defined as  $T_m$  and  $T_v$ .

3) *Derivation of Features*: Basic features of a blinking pattern are made up of 15 features in the time domain and 6 features in frequency domain. However, blinks vary from 10 to 36 times in each segment of our dataset so that we select the maximum, minimum and average values of each feature in the time domain to extend 15 features to 45. Until now, 45 features in time domain and 6 features in frequency domain are extracted from a single curve containing a certain number of blinks. In addition, when processing the event density, we generate 4 curves from 4 different types of density streams, namely, a positive event density stream (ON curve), negative eye blink stream (OFF curve), spliced eye blink stream (Splice curve), and overlapped eye blink stream (Overlap curve). Therefore, 51 features are quadrupled to form a feature set of 204 candidates. Visualization of the 4 types of event stream can be found in Figs. 6(a) and 6(b). It is obvious that some of features are highly correlated. Therefore, it is necessary to conduct some experiments about feature selection to improve the feature sets.

#### IV. NEUROBIOMETRIC FEATURE SELECTION

To determine the most significant features, we apply recursive feature elimination (RFE) on the candidate set to produce feature importance rankings and a reduced feature subset. Then we further compress the subset size by two different approaches, one with the Pearson correlation coefficients based approach and the other using a coefficient of variation ( $CoV$ ) based bi-variate analysis.

##### A. Recursive Feature Elimination-Based Approach

First we calculate a feature weights ranking with the recursive feature elimination method, using an estimator (e.g., SVM) that assigns weights to each candidate feature. We recursively consider smaller and smaller subsets of features by pruning the least important one until the desired size is reached. During each iteration, the estimator's performance over the subset can be concluded by re-training and re-evaluating several times to ensure stability. Then these test points are fitted by a polynomial trend curve and the size of its extreme point is taken as the new subset size. By doing so, we select 135 significant features out of 204 candidates.

##### B. Pearson Correlation Coefficient-Based Approach

To further compress the subset size from step one, we

conduct the Pearson correlation coefficient based method by calculating the Pearson correlation coefficient and its  $p$ -value between each pair of features. In the matrix of feature correlations, weaker correlation means they are more independently distributed. Thus feature pairs with a correlation coefficient higher than  $r$  and with a  $p$ -value below  $\alpha$  are selected. Then, we sort features by the number of times each feature been selected in decreasing order. Lastly, the top  $x_p$  percent features are removed.

##### C. Coefficient of Variation-Based Approach

Features with more variation across subjects have a better chance of capturing subjects' distinctive information. And  $CoV$  can be a good measure of this variation. First, we calculate the mean  $\mu_j$  and standard deviation  $\sigma_j$  for every feature  $j$  across subjects, and compute  $CoV_j = \sigma_j/\mu_j$ . Then, we rank all features in descending order of their  $CoV$  values computed across all subjects. In addition, we compute the range  $R_i^j = [\mu_i^j - \sigma_i^j, \mu_i^j + \sigma_i^j]$  for each feature on each subject, after which we count the number of subject- $k$  whose  $\mu_i^k$  lies within range  $R_i^j$ . The count reflects the number of subjects that share distribution on this feature and may lead to misidentification. Thus, features with lower counts are better in distinguishing subject pairs. We rank features in an increasing order of their counts. Finally, we average the above two rankings and only preserve a certain proportion from the list bottom, removing a  $x_c$  percent of features.

## V. IDENTIFICATION AND VERIFICATION

### A. Identification Mode

To evaluate the feature selection methods and the two classifiers in identification mode, 10% of the dataset for each subject is selected to generate the test samples, the identity of each testing subject is generated by classifiers. The  $ACC$  for this experiment represents the fraction of subject identity predictions that are correct. The  $ACC$  is calculated as follows:

$$ACC(i) = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where  $TP$ ,  $TN$ ,  $FP$ ,  $FN$  stand for true positives, true negatives, false positives, and false negative, respectively. After 10 times repeat, the average  $ACC$  can be calculated as follows:

$$ACC_{av} = \frac{1}{10} \sum_{i=1}^{10} ACC(i). \quad (2)$$

Besides, we evaluate the methods with a new variable, feature remain ratio ( $FRR$ ), to measure the remaining feature number that a feature selection approach preserves. A lower  $FRR$  value means that the system requires fewer features for authentication. Hence, it is considered with better efficiency and usability.

$$FRR = \frac{F_{\text{selected}}}{F_{\text{total}}}. \quad (3)$$

In identification mode,  $ACC$  is evaluated under a number of features that are selected with the principle described in Section III. It is obvious that the higher  $ACC$  and lower  $FRR$

are, the better the system is.

### B. Verification Mode

To evaluate the feature selection methods and the classifiers in verification mode, each subject in the database tries to access the system with the true identity (the genuine identity in the database) and false identity (one of the remaining identities in the database), respectively. The evaluation criteria are listed as follows: false positive rate (*FPR*) defines the possibility of an imposter accepted by the proposed system:

$$FPR = \frac{FP}{TN + FP}. \quad (4)$$

False negative rate (*FNR*) defines the fraction of a genuine subject rejected by the proposed system

$$FNR = \frac{FN}{FN + TP}. \quad (5)$$

Obviously, an ideal system should have both lower *FPR* and *FNR*, but higher *ACC*. In verification mode, *FPR* and *FNR* are evaluated under different number of features that are selected obeying the principle described in Section III.

### C. Authentication Models

1) *Authentication With Support Vector Machine*: Support vector machine is a supervised classifier that is widely used.

Its kernel maps data from low-dimension space to higher-dimension space. Different kernel functions results in very different prediction results. We test over four most commonly used kernel functions: linear kernel, polynomial kernel, RBF kernel, and sigmoid kernel. Results show that a linear kernel SVM with a squared penalty parameter of 0.75, a squared hinge loss and a one-versus-one multi-class strategy could achieve the best performance among non-ensemble models. Thus, this linear SVM is chosen for later experiments.

2) *Authentication With Bagging Linear Discriminant Analysis*: The algorithm uses a set of LDA models as base-classifiers to build an ensemble classifier. The LDA model is a classifier with a linear decision boundary, generated by using Bayes' rule. It fits a Gaussian density to each class. Then each base classifier fits random subsets of the dataset using the bagging (bootstrap aggregating) method and aggregates their predictions, to obtain the final prediction. It reduces the variance of base estimators through randomization and aggregation. In the bootstrap sampling, we sample 80% of samples and 100% of features from the total train set. Then  $k$  LDA models are fitted using singular value decomposition accordingly. The final output identity of each sample is determined by voting aggregation. Though increasing the number of  $k$  could enhance the model's fitting capability at the cost of inflating time consumption, in our experiments, 30 is taken as a good compromise of computation speed and accuracy.

## VI. EXPERIMENTS

### A. Parameter Sets

In this section, we mainly discuss how to find the optimal parameter set for each procedure in the proposed algorithm.

Table II summarizes the relative parameters.

TABLE II  
SUMMARY OF HYPER PARAMETERS IN THE PROPOSED ALGORITHM

Procedure	Symbol	Description
Filtration and curve smoothing	$N$	Number of adjacent pixels to communicate
	$t$	Threshold to determine the validity of the central pixel
	$w$	Smooth width
Feature extraction	$N_b$	Number of blinks to merge as one
	$s_b$	Scale of the starting threshold during a blink
Feature selection	$k_{RFE}$	Number of features selected by RFE
	$x_p$	Percent of redundant features removed by P90
	$r$	Correlation threshold, $r \in \{0.9, 0.95\}$
	$\alpha$	Statistical significance level, we use $\alpha = 0.05$
	$x_c$	Percent of redundant features removed by $CoV$

1) *Filtration and Curve Smoothing*: The hyper parameters related to filtration and curve smoothing are the number of adjacent pixels to communicate  $N$ , the threshold to determine the validity of the central pixel  $t$ , and the smooth width  $w$ . Only the facial region, including the eyebrow and ocular areas, is recorded in the lens area. We regard the lens area as a two-dimensional tensor. The areas where the event signals are triggered by the eye blink are concentrated and large, while the areas formed by the noise triggered event signals are discrete and small. We test three odd values (3, 5, 7) for  $N$  and consider the central pixel whose adjacent pixels account for more than 1/3 of the area as a valid one. Results show that a  $5 \times 5$  area with the threshold equaling to 8 could achieve the best performance in the filtration method. The time unit of each frame is 2 ms and the smooth width is 5 in our work.

2) *Feature Extraction*: As for the number of blinks  $N_b$ , we average the features extracted from  $n$  neighboring eye blinks respectively and train the mean with the proposed linear SVM. The number of blinks ranges from 1 to 7. In Fig. 8(c), we observe that with an increase of blinks in a trial, the values of *FNR* and *FPR* draws nearer, with *FPR* largely remaining steady and *FNR* gradually decreasing. At the same time, *ACC* climbs as the result of more abundant user information. The variations gradually become non-significant when blink number reaches 5 and above. Though more blinks bring more subject characteristics for successful authentication, it becomes more inconvenient and annoying for the users to complete a trial. Therefore, a compromise is made to set the number blinks to 5 in a single trial. An example showing 5 blinks and its fitting curve can be found in Figs. 9(a) and 9(b). When it comes to scale of the starting threshold of a blink  $s_b$ , we observe that the maximum peak of eye closing is about 100 times the start value of eye closing (shown in Fig. 4) in 4 curves. Therefore, we set  $s_b$  to 0.01.

3) *Feature Selection*: To obtain the optimal parameter set for each feature selection method, we first calculate the average accuracy and *FRR* values from all these subsets. Statistical significance level  $\alpha$ , the primary parameter of



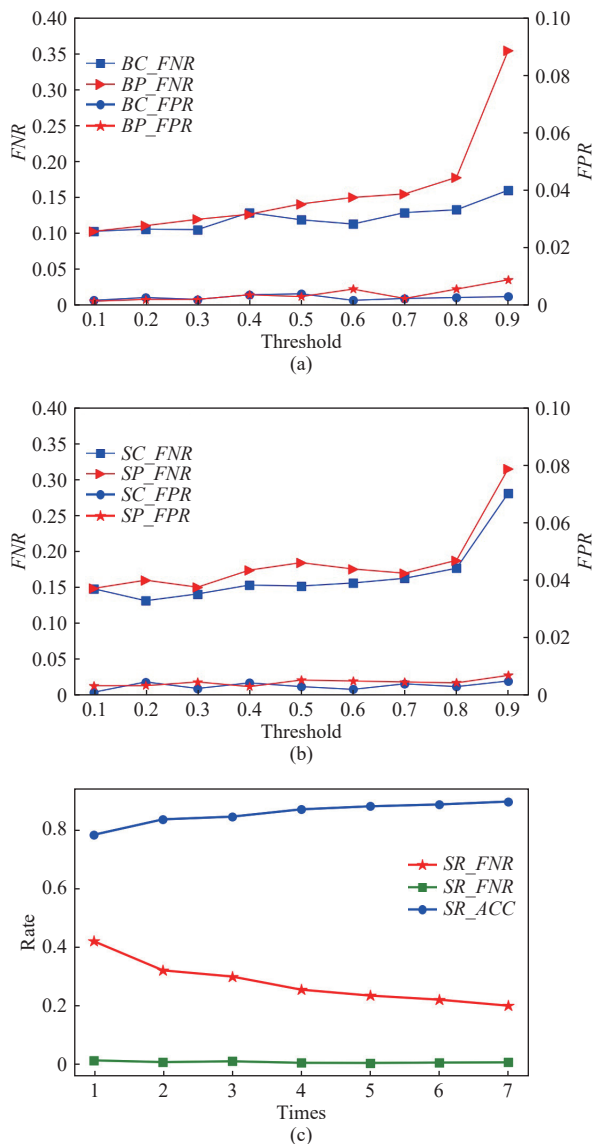


Fig. 8. (a)  $FNR$  and  $FPR$  across different thresholds for bagging linear discriminant.  $BC$  indicates that the feature sets are selected by  $CoV$  and  $BP$  indicates that the feature sets selected by the Pearson correlation coefficient. (b)  $FNR$  and  $FPR$  across different thresholds for SVM. (c)  $ACC$ ,  $FNR$ , and  $FPR$  across different numbers of eye blinks in a trial as in Fig. 9(a) demonstrated using SVM.  $SR$  means that the feature sets are selected only by RFE approach.

Pearson correlation coefficient-based method, determines whether the correlation is obvious or not. In this experiment, we set  $\alpha$  as 0.05. The correlation is measured by correlation threshold  $r$ , whose positive value is the positive correlation among variables. Threshold  $r$  varies from 0.9 to 0.95, with an increase of 0.01 in each trial. However, we observe that the difference is insignificant among our settings so that we set  $r$  to 0.9. On top of that, we perform the graphical parameter selection approach, by plotting the correlation graph between thresholds and the above metrics as shown in Fig. 10.

We can observe that as the thresholds of  $x_p$  or  $x_c$  increase, the average accuracy and  $FRR$  gradually decrease. And as  $FRR$  falls, the accuracy remains largely unaffected until  $FRR$

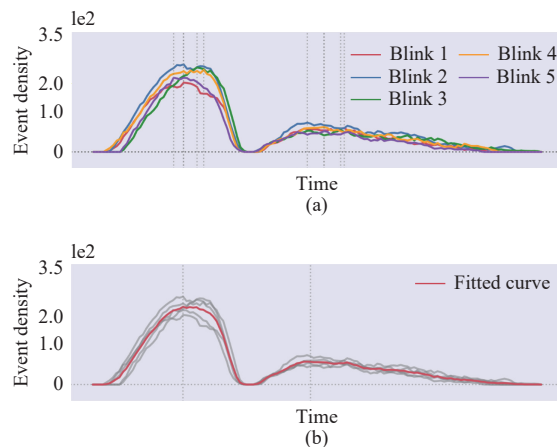


Fig. 9. The visualization of event density curves for 5 eye blinks and their median fitting curve (better viewed in color).

goes below 10% which is when the average accuracy starts noticeably dropping. Thresholds  $x_p$  and  $x_c$  of 0.1 achieves the best average accuracy in most scenarios and thus be chosen in our evaluation, but when compromising for the less amount of feature numbers (i.e., lower  $FRR$ ), we could even push the threshold up to 0.8, which means keeping only 20 percent from the already reduced feature set, and still maintaining the accuracy steady.

### B. Cross Validation Versus Random Segmentation

To verify the rationality of the data partition, we perform 10-fold cross-validation and random segmentation (RS) on the data set, respectively. As for 10-fold cross-validation, two methods are adopted to split the data set, one of which (named 10-fold-A) guarantees the consistency of the sample numbers for each subject in both the training and testing sets while the other (named 10-fold-B) selects 10% of the whole data set as test samples. As for RS, each subject has the same number of samples in each set. Of the features selected by  $CoV$ , all the generated sets are trained with SVM. Results show that RS reaches the lowest accuracy of the three methods, around 84%, while those split by 10-fold-B gain the highest accuracy, about 92%. Meanwhile, the contingency of RS makes models inaccurate. When considering the sample equilibrium, 10-fold-A with the accuracy of about 87% is chosen as the final partition method.

### C. Ensemble Versus Non-Ensemble Authentication

According to the result, we can determine that the proposed system, i.e., SVM with a linear kernel function and the LDA model, can reject almost all imposters and reach around 87% accuracy when evaluating acceptance of the genuine users. On average, each user can pass the authentication of the system within 0.3 times of trial. Fig. 10 shows that the features selected by  $CoV$  gain a higher accuracy in both the identification and verification model than Pearson selection in most cases.

Table III displays the authentication summary with both the ensemble model and non-ensemble model. Both models test the feature list selected by RFE,  $CoV$ , and P90, where P90

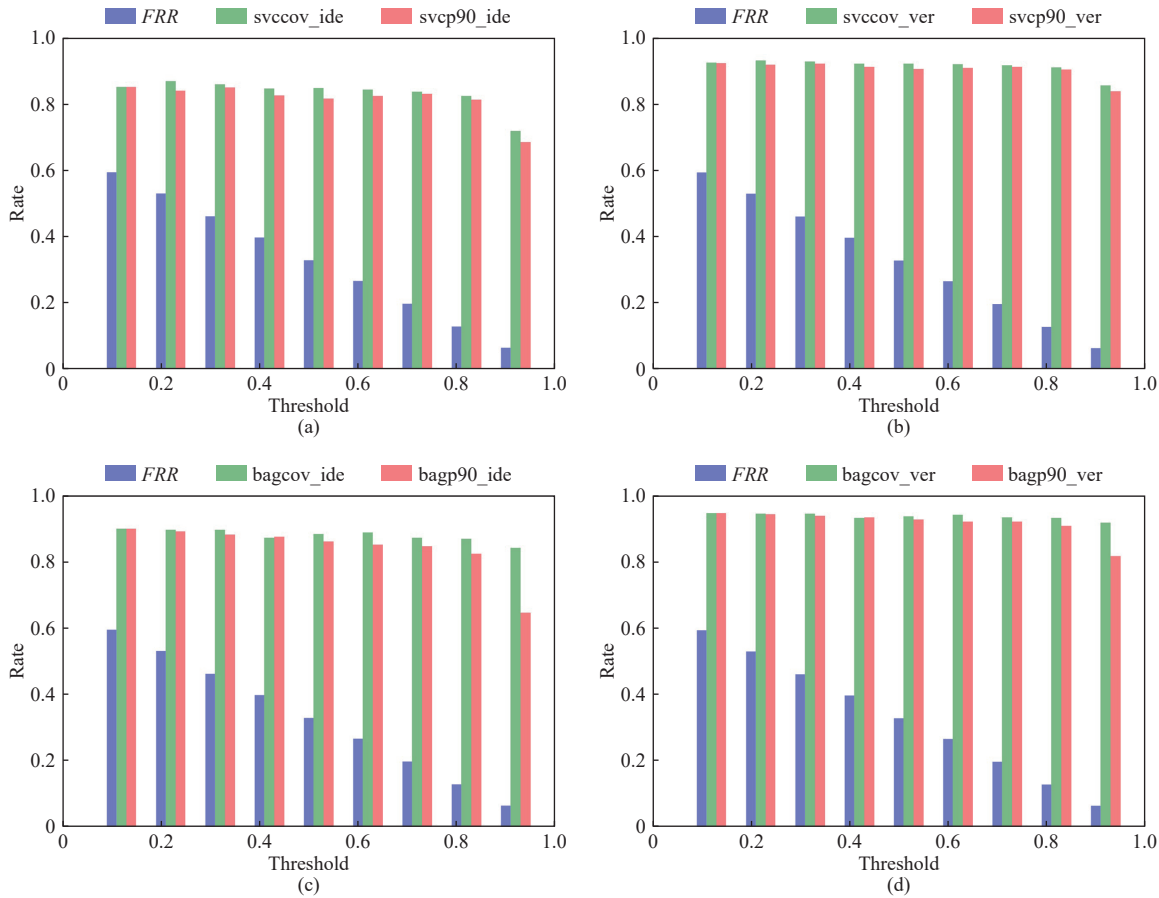


Fig. 10. (a) *FRR* and *ACC* across different thresholds for identification using SVM. *CoV* refers to the coefficient of variation based feature selection approach. *P90* refers to the Pearson correlation coefficient based feature selection approach with a threshold  $r$  of 0.9. (b) *FRR* and *ACC* across different thresholds for verification using SVM. (c) *FRR* and *ACC* across different thresholds for identification using bagging linear discriminant. (d) *FRR* and *ACC* across different thresholds for verification using bagging linear discriminant (better viewed in color).

TABLE III  
 AUTHENTICATION SUMMARY OF FEATURE SELECTION APPROACHES AND AUTHENTICATION MODELS. SYMBOL MEANINGS:  $n$  FOR SELECTED FEATURE SET SIZE,  $N$  FOR THE ORIGINAL FEATURE SET SIZE,  $W$  FOR NUMBER OF BLINKS IN A TRIAL

Method	Feature selection			Ensemble model			Non-ensemble model		
	$n$	$N$	$W$	<i>ACC</i>	<i>FNR</i>	<i>FPR</i>	<i>ACC</i>	<i>FNR</i>	<i>FPR</i>
RFE	135	204	5	0.9498	0.0987	0.0016	0.9263	0.0987	0.0016
<i>CoV</i>	121	135	5	0.9479	0.1026	0.0016	0.9257	0.1477	0.0010
<i>P90</i>	121	135	5	0.9481	0.1026	0.0013	0.9242	0.1484	0.0033

refers to the Pearson correlation coefficient based feature selection approach with a threshold  $r$  of 0.9. The authentication context based on the selection method *CoV* and *P90* can be more concise because the feature set size used by them is already reduced by RFE. They also share the same feature size at the same time. Furthermore, the feature sets selected by *CoV* and *P90* are similar to each other. This proves both the selection methods are effective to some degree. From Table III, we can determine that the ensemble model reaches a higher *ACC* and lower *FNR* compared the other models and that the *P90* based method can select better features than *CoV* in the selection step 2 of ensemble model. The system with ensemble model and *P90* selection method presents a higher availability and security ( $ACC \approx 0.9481$ ,  $FNR \approx 0.1026$ ,  $FPR \approx 0.0013$ ).

D. Error Analysis

In this section, we mainly evaluate *CoV* and *P90* methods, where Table III shows that the RFE-only method has similar authentication performance but requires more features to train with.

We discuss the impact on verification brought by altering thresholds in Figs. 8(a) and 8(b).

Both figures present different performance measures in varied selection methods with the threshold varying from 0.1 to 0.9. The threshold increases 0.1 at each step. We introduce *FPR* and *FNR* to represent the possibility of an imposter (with an arbitrary false identity included in the dataset) accepted by the system and a genuine subject rejected by it. All the subjects (45 in total) share the same ensemble

and non-ensemble models which are trained to identify the genuine user. The fraction of imposter acceptance is nearly  $1/45$  of that of genuine user rejection. So  $TN$  is higher than  $TP$ , leading to lower  $FPR$ . Even though  $FPR$  is much lower than  $FNR$  and  $EER$  (the point where  $FNR$  equals to  $FPR$ ) may be off the chart,  $FPR$  and  $FNR$  are both below 0.2 when the threshold is less than 0.8. In addition, our experiments show that each user can go through authentication within 0.3 times of the trial. Therefore, the system can be accepted.  $FNR$  and  $FPR$  of both models have similar values, keeping in line with the threshold varying from 0.1 to 0.8 and increasing rapidly (except the  $FNR$  of ensemble model with the  $CoV$  method) when the system removes 90% features. In addition, Fig. 10 shows decreased accuracy with a threshold of 0.9. This can be explained by the fact that the remaining feature sets are too small to depict the subject precisely. In Fig. 8(a), we can observe that the ensemble model using the threshold of 0.7 can gain a lower  $FPR$  at the cost of a small increase in  $FNR$ , which would be a good choice for faster computation and future utility. In terms of the non-ensemble model, it can produce a desirable result with a threshold of 0.6 ( $FNR \approx 0.15$ ,  $FPR \approx 0.002$ ).

#### E. Number of Subjects

Since there is no other open data set to evaluate our proposed algorithm, we reduce the number of subjects to repeat the experiments. The chosen system is the ensemble model and P90 selection method (with a threshold  $x_p$  equals to 0.3), which is proved to be the most effective method in Sections VI-C and VI-D. The number of subjects varies from 20 to 45, with an increase of 5 in each step. Fig. 11 shows that as the number of subjects increases, both the identification and verification accuracy remain largely unchanged while  $FPR$  gradually decreases. The test with 40 subjects achieves the best average identification and verification accuracy. The more subjects participate in the training, the lower  $FPR$  is. Therefore, our authentication system demonstrates both usability and security.

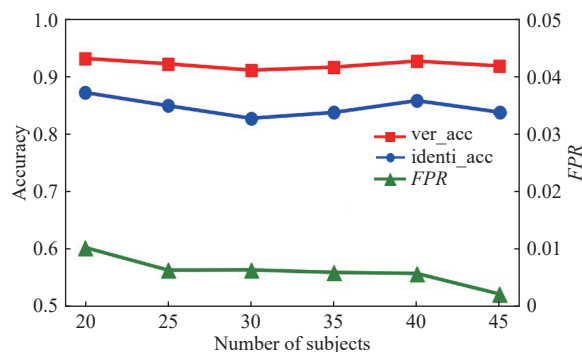


Fig. 11. Accuracy and  $FPR$  across different numbers of subjects for identification and verification using bagging linear discriminant.

## VII. DISCUSSION

#### A. Event Density Biometric

The results show that using the sole biometric of event

density can still enable reliable identity authentication with promising performance. This type of authentication system is very easily equipped, as it only requires one dynamic vision sensor to capture natural eye blinks from a subject.

#### B. Usability and Security to Blink Times

The system's usability and security level is a compromise that can be adjusted according to different users' needs or application scenarios. By increasing the number of blinks in a single trial, as  $ACC$  increases and  $FNR$  decreases, the security level is generally improved. However, the need for more blinks adds to the unwillingness to perform such authentication and thus reduces the system's overall usability. Fortunately, the critical  $FPR$  is reduced to a very low level and is not significantly changed with different blink numbers. Therefore, though we have discovered that a blink number of 5 is sufficient and efficient enough in the trade-off, a smaller number of blinks can also provide enough security guarantee with a low  $FPR$  that rejects most false identity claims.

#### C. System Robustness

The system is proved to be capable of rejecting false claims in our experiments with low  $FPR$  and thus can be considered robust against an identity attack from a living human. As for facing synthetic data that resembles a living human, our authentication system is still able to function because it is built using a particular biometric, the event density, which comes with an extremely high scanning rate (e.g., 500 Hz in our implementation, and can up to 1000 Hz) and low response latency. These advantages result in outstanding temporal resolution and enable the system to capture and learn more precise details from tested data. Moreover, the use of multiple event types (ON, OFF, Overlay, and Splice event curves) further consolidates the robustness of the authentication. However, discussing the performance of living subject detection is not a focus of this paper, and we may investigate it in other works.

#### D. Face Movement

Current works only record and process the events triggered by eye blinking under conditions where the face is still. However, blinking, an unconscious movement triggered by biological currents, is more biologically meaningful in the field of authentication in nature and more individually encrypted compared to other explicit biometrics. Research based on simple eye movements can better extract essential information. What's more, judging from the current state of face recognition systems, users are comfortable with the requirement to keep their faces motionless when using the systems. As for synthetic movement that is overlapped by face and eye movements, we proposed that eye movement can be separated utilizing motion compensation or priori knowledge. We will investigate this in future works.

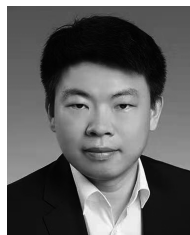
## VIII. CONCLUSION

Our work introduces the first-ever neuromorphic sensor-based human authentication system using the easy-capture biometrics of eye blinks. One neuromorphic sensor such as a

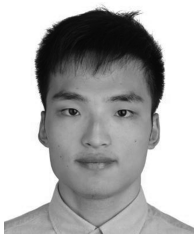
dynamic vision sensor can be used to accurately identify and verify users with relatively simple computational processes. The detailed results show that the system is effective and efficient based on different feature selection approaches and authentication models. The system only uses implicit and highly dynamic features of the user's eye blinks at a microsecond level time resolution and at an asynchronous pixel-level vision resolution, which not only ensures security but also avoids recording visible characteristics of a user's appearance. This work demonstrates a new way towards safer authentication using neuromorphic vision sensors. In future works, experiments may extend into the effects of the temporal evolution of human growth or other changes. Also, to improve the system's robustness against attacks, adversarial samples may be added in the future to provide more safety margin.

## REFERENCES

- [1] M. M. Ali, V. H. Mahale, P. Yannawar, and A. Gaikwad, "Overview of fingerprint recognition system," in *Proc. IEEE Int. Conf. Electrical, Electronics, and Optimization Techniques*, pp. 1334–1338, 2016.
- [2] Y. Zhang and M. Juhola, "On biometrics with eye movements," *IEEE J. Biomedical and Health Informatics*, vol. 21, no. 5, pp. 1360–1366, 2016.
- [3] Z. Li, Y. Yuan, L. Luo, W. Su, K. Zhao, C. Xu, J. Huang, and M. Pi, "Hybrid brain/muscle signals powered wearable walking exoskeleton enhancing motor ability in climbing stairs activity," *IEEE Trans. Medical Robotics and Bionics*, vol. 1, no. 4, pp. 218–227, 2019.
- [4] Z. Li, J. Li, S. Zhao, Y. Yuan, Y. Kang, and C. P. Chen, "Adaptive neural control of a kinematically redundant exoskeleton robot using brain-machine interfaces," *IEEE Trans. Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3558–3571, 2018.
- [5] Y. Liu, W. Su, Z. Li, G. Shi, X. Chu, Y. Kang, and W. Shang, "Motor-imagery-based teleoperation of a dual-arm robot performing manipulation tasks," *IEEE Trans. Cognitive and Developmental Systems*, vol. 11, no. 3, pp. 414–424, 2018.
- [6] Q. Wu, Y. Zeng, C. Zhang, L. Tong, and B. Yan, "An EEG-based person authentication system with open-set capability combining eye blinking signals," *Sensors*, vol. 18, no. 2, pp. 335, 2018.
- [7] C. Gou, Y. Wu, K. Wang, K. Wang, F.-Y. Wang, and Q. Ji, "A joint cascaded framework for simultaneous eye detection and eye state estimation," *Pattern Recognition*, vol. 67, pp. 23–31, 2017.
- [8] M. Jeong and B. C. Ko, "Drivers facial expression recognition in real-time for safe driving," *Sensors*, vol. 18, no. 2, pp. 4270, 2018.
- [9] V. K. Sinha, A. K. Gupta, and M. Mahajan, "Detecting fake iris in iris bio-metric system," *Digital Investigation*, vol. 25, pp. 97–104, 2018.
- [10] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A novel biometric approach for human identification and verification using eye blinking signal," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 876–880, 2014.
- [11] J. A. Lenero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, "A 3.6  $\mu$ s latency asynchronous frame-free event-driven dynamic-vision sensor," *IEEE J. Solid-State Circuits*, vol. 46, pp. 1443–1455, Jun. 2011.
- [12] Z. Sitova, J. Sedenka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani, "HMOG: New behavioral biometric features for continuous authentication of smart phone users," *IEEE Trans. Information Forensics and Security*, vol. 11, May 2016.
- [13] N. Merhav, "False-accept/false-reject trade-offs for ensembles of biometric authentication systems," *IEEE Trans. Information Theory*, vol. 65, no. 8, Aug. 2019.
- [14] D. Q. Phung, D. Tran, W. Ma, P. Nguyen, and T. Pham, "Using shannon entropy as EEG signal feature for fast person identification," in *Proc. 22nd European Symp. Artificial Neural Networks, Computational Intelligence and Machine Learning*, vol. 4, pp. 413–418, 2014.
- [15] H. U. Amin, A. S. Malik, R. F. Ahmad, N. Badruddin, N. Kamel, M. Hussain, and W.-T. Chooi, "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Australasian Physical & Engineering Sciences in Medicine*, vol. 38, no. 1, pp. 139–149, 2015.
- [16] Y. Ding, X. Hu, Z. Xia, Y.-J. Liu, and D. Zhang, "Inter-brain EEG feature extraction and analysis for continuous implicit emotion tagging during video watching," *IEEE Trans. Affective Computing*, 2018. DOI: 10.1109/TAFFC.2018.2849758
- [17] S. N. Abbas and M. Abo-Zahhad, "Eye blinking EOG signals as biometrics," in *Biometric Security and Privacy*, pp. 121–140, Springer, 2017.
- [18] S. Yang, S. Hoque, and F. Deravi, "Improved timefrequency features and electrode placement for eegbased biometric person recognition," *IEEE Access*, vol. 7, pp. 49604–49613, 2019.
- [19] A. George and A. Routray, "Fast and accurate algorithm for eye localisation for gaze tracking in lowresolution images," *IET Computer Vision*, vol. 10, no. 7, pp. 660–669, 2016.
- [20] I. Rigas and O. V. Komogortsev, "Biometric recognition via probabilistic spatial projection of eye movement trajectories in dynamic visual environments," *IEEE Trans. Information Forensics and Security*, vol. 9, no. 10, pp. 1743–1754, 2014.
- [21] Z. Wang, C. Li, H. Shao, and J. Sun, "Eye recognition with mixed convolutional and residual network (micore-net)," *IEEE Access*, vol. 6, pp. 17905–17912, 2018.
- [22] J. Espinosa, B. Domenech, C. Vázquez, J. Pérez, and D. Mas, "Blinking characterization from high speed video records. Application to biometric authentication," *PLoS one*, vol. 13, no. 5, pp. e0196125, 2018.
- [23] S. Seha, G. Papangelakis, D. Hatzinakos, A. S. Zandi, and F. J. Comeau, "Improving eye movement biometrics using remote registration of eye blinking patterns," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pp. 2562–2566, 2019.
- [24] G. Chen, L. Hong, J. Dong, P. Liu, J. Conradt, and A. Knoll, "EDDD: Event-based drowsiness driving detection through facial motion analysis with neuromorphic vision sensor," *IEEE Sensors J.*, vol. 20, no. 11, pp. 6170–6181, 2020.
- [25] G. Chen, W. Chen, Q. Yang, Z. Xu, L. Yang, J. Conradt, and A. Knoll, "A novel visible light positioning system with event-based neuromorphic vision sensor," *IEEE Sensors J.*, vol. 20, no. 17, pp. 10211–10219, 2020.
- [26] A. Yousefzadeh, G. Orchard, T. Serrano-Gotarredona, and B. Linares-Barranco, "Active perception with dynamic vision sensors. Minimum saccades with optimum recognition," *IEEE Trans. Biomedical Circuits and Systems*, vol. 12, no. 4, pp. 927–939, 2018.
- [27] G. Lenz, S.-H. Ieng, and R. Benosman, "High speed event-based face detection and tracking in the blink of an eye," arXiv preprint arXiv: 1803.10106, 2018.
- [28] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza, "Event-based vision: A survey," arXiv: 1904.08405, 2019. DOI: 10.1109/TPAMI.2020.3008413.
- [29] G. Chen, H. Cao, J. Conradt, H. Tang, F. Röhrbein, and A. Knoll, "Event-based neuromorphic vision for autonomous driving: A paradigm shift for bio-inspired visual sensing and perception," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 34–49, 2020.



**Guang Chen** (M'19) is a Research Professor at Tongji University and a Senior Research Associate (guest) at Technical University of Munich, Germany. He is leading the Intelligent Perception System Group at Tongji University. His research interests include computer vision, image processing and machine learning, and the bio-inspired vision with applications in robotics and autonomous vehicle. He was a Research Scientist at Fortiss GmbH, a research institute of Technical University of Munich from 2012 to 2016, and a Senior Researcher at the Chair of Robotics, Artificial Intelligence and Real-time Systems, Technical University of Munich from 2016 to 2017. He received the Ph.D degree in the Faculty of Informatics in Technical University of Munich. He received the B.S and M.Eng degrees in mechanical engineering from Hunan University.



**Fa Wang** received the B.E. degree and is currently pursuing M.E. degree at the School of Automotive Studies of Tongji University. His research interests include computer vision, deep learning, lidar perception, and autonomous driving.



**Xiaoding Yuan** is currently pursuing the B.E. degree in the College of Electronics and Information Engineering of Tongji University. Her research spans the areas of computer vision and machine learning.



**Zhijun Li** (M'07–SM'09) received the Ph.D. degree in mechatronics from Shanghai Jiao Tong University in 2002. From 2003 to 2005, he was a Postdoctoral Fellow in Department of Mechanical Engineering and Intelligent systems, The University of Electro-Communications, Japan. From 2005 to 2006, he was a Research Fellow in the Department of Electrical and Computer Engineering, National University of Singapore, and Nanyang Technological University, Singapore. From 2017, he is a Professor in Department of Automation, University of Science and Technology. From 2019, he is the Vice Dean of School of Information Science and Technology, University of Science and Technology of China. From 2016, he has been the Co-Chairs of IEEE SMC Technical Committee on Bio-mechatronics and Bio-

robotics Systems (B2S), and IEEE RAS Technical Committee on Neuro-Robotics Systems. He is serving as an Editor-at-large of Journal of Intelligent & Robotic Systems, and Associate Editors of several IEEE Transactions. Dr. Li's current research interests include wearable robotics, tele-operation systems, nonlinear control, neural network optimization, etc.



**Zichen Liang** received the B.S. degree and is currently pursuing the M.E. degree at the School of Automotive Studies of Tongji University. Her current research interests include deep learning and computer vision.



**Alois Knoll** received the diploma (M.Sc.) degree in electrical/communications engineering from the University of Stuttgart, Germany, in 1985 and the Ph.D. (summa cum laude) in computer science from the Technical University of Berlin, Germany, in 1988. He served on the Faculty of the Computer Science Department of TU Berlin until 1993. He joined the University of Bielefeld, as a Full Professor and the Director of the Research Group Technical Informatics until 2001. Since 2001 he has been a Professor at the Department of Informatics, TU Munich. He was also on the board of Directors of the Central Institute of Medical Technology at TUM (IMETUM). From 2004 to 2006, he was Executive Director of the Institute of Computer Science at TUM. His research interests include cognitive, medical and sensor-based robotics, multi-agent systems, data fusion, adaptive systems, multimedia information retrieval, model-driven development of embedded systems with applications to automotive software and electric transportation, as well as simulation systems for robotics and traffic.