Quantifying Limb Movements in Epileptic Seizures Through Color-Based Video Analysis

Haiping Lu*, *Member, IEEE*, Yaozhang Pan, *Member, IEEE*, Bappaditya Mandal, How-Lung Eng, *Member, IEEE*, Cuntai Guan, *Senior Member, IEEE*, and Derrick W. S. Chan

***Abstract*—This paper proposes a color-based video analytic system for quantifying limb movements in epileptic seizure monitoring. The system utilizes colored pyjamas to facilitate limb segmentation and tracking. Thus, it is unobtrusive and requires no sensor/marker attached to patient's body. We employ Gaussian mixture models in background/foreground modeling and detect limbs through a coarse-to-fine paradigm with graph-cut-based segmentation. Next, we estimate limb parameters with domain knowledge guidance and extract displacement and oscillation features from movement trajectories for seizure detection/analysis. We report studies on sequences captured in an epilepsy monitoring unit. Experimental evaluations show that the proposed system has achieved comparable performance to EEG-based systems in detecting motor seizures.**

***Index Terms*—Color-based, epilepsy, monitoring, movement quantification, seizure, video.**

I. Introduction

EPILEPTIC seizures are transient symptoms of epilepsy, a serious brain dynamical disorder, due to abnormal excessive or synchronous neuronal activity in the brain [1]–[3]. A popular technique for seizure detection is electroencephalogram (EEG) [4]–[9]. For example, the work in [7] extracts features for seizure recognition from EEG signals arranged as a third-order tensor [10], [11]. EEG-based approach requires patients to wear electrodes on scalp and remain attached to EEG equipment during monitoring.

Many epileptic seizures induce uncoordinated movements in a patient's body, such as jerking or stiffening. Unfortunately, traditional naked-eye examination does not allow for quantitative analysis of the amplitude and frequency of movements in seizures [12]. Some researchers attempt to detect seizures through quantifying body movements with attached markers or sensors [12]–[14]. The works in [12] and [13] attach infrared reflective markers or white foam markers to a patient's body for movement quantification. However, these markers are often identified manually with mouse clicking and lead to correspondence difficulties. Thus, only short sequences can be analyzed. For example, the results in [13] show only up to 23 s, while only about 14 s is analyzed in [12]. On the other hand, the work in [14] performs visual analysis of signals from accelerometers worn on a patient's body to detect seizures. As attached markers/sensors often lead to discomfort and are susceptible to dislocation over time, these solutions are problematic for long-term monitoring.

This paper focuses on unobtrusive video-based seizure detection. To the best of our knowledge, the only existing system in this category is developed by Karayiannis *et al.* for neonates [15]–[19]. They select anatomic sites on moving limbs by thresholding the motion vector magnitudes and then track the selected sites. They quantify limb motion by temporal motion-strength signals extracted from video segments. As this approach relies purely on motion information, the level of quantification details is limited and coarse. Furthermore, as a motion-based method, tracking tends to become unreliable over time. Thus, the maximum length of the processed video segments in [15]–[19] is only 20 s and this approach is not suitable for long-term monitoring either.

This paper proposes a new unobtrusive color-based video analytic system for quantifying limb movements in epileptic seizure monitoring, as shown in Fig. 1 (a preliminary version of this work was reported in [20]). We aim to quantify movements in greater details than the motion-based system in [15]–[19] for sequences of longer durations. The proposed system requires a patient to wear pyjamas with limb portion colored, instead of attaching sensors or markers. A camera mounted on the ceiling captures videos of patient activities in a clinical epilepsy monitoring unit (EMU). We excerpt epochs of activities and perform automated video analysis following a simple manual initialization. Specifically, we detect the limbs, quantify them with three parameters each, and extract features from parameter trajectories for seizure detection/analysis.

The main contributions of this study are as follows.

1) Introduction of a color-pyjamas-based video analytic system for patient monitoring. It does not attach anything to the body, providing more comfortable experience than those in [7] and [12]–[14]. Furthermore, colored pyjamas can be worn easily. In contrast, electrodes, markers, or sensors often need a professional to attach to a patient.
2) Proposal of limb movement quantification and seizure detection algorithms for the color-based video system. Besides a simple user initialization on the first frame, the

Manuscript received July 18, 2012; revised October 14, 2012; accepted October 22, 2012. Date of publication November 21, 2012; date of current version January 16, 2013. This work was supported by the Science and Engineering Research Council, Agency for Science, Technology and Research, Singapore. *Asterisk indicates corresponding author.*

*H. Lu is with the Institute for Infocomm Research, Agency for Science, Technology and Research, 138632 Singapore (e-mail: hplu@ieee.org).

Y. Pan, B. Mandal, H.-L. Eng, and C. Guan are with the Institute for Infocomm Research, Agency for Science, Technology and Research, 138632 Singapore (e-mail: yzpan@i2r.a-star.edu.sg; bmandal@i2r.a-star.edu.sg; hleng@i2r.a-star.edu.sg; ctguan@i2r.a-star.edu.sg).

D. W. S. Chan is with KK Women's and Children's Hospital, 138632 Singapore (e-mail: derrick.chan.ws@kkh.com.sg).

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/TBME.2012.2228649

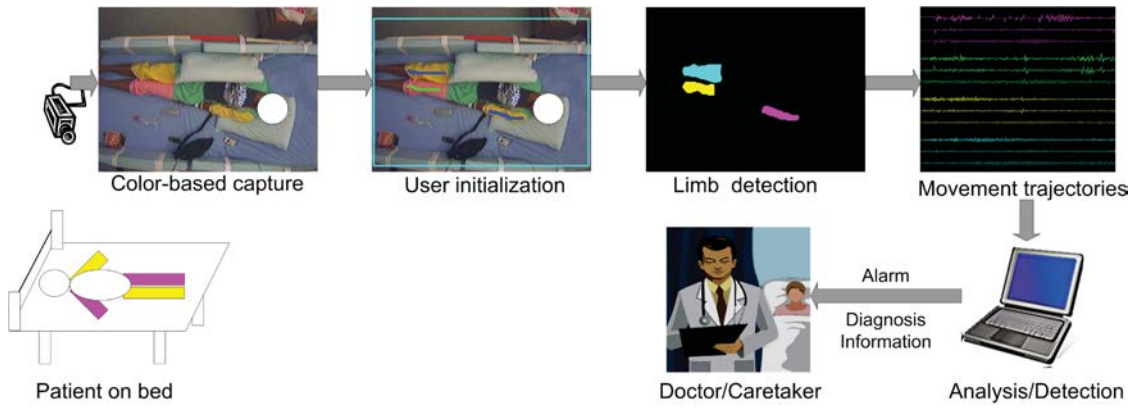

Fig. 1. Illustration of the proposed video analytic system. Limb movements are detected based on color pyjamas and then analyzed for epileptic seizure detection. The patient's face is masked in this paper for anonymity and this is not part of the user initialization.

algorithms are fully automatic and require no manual input, unlike [12] and [13]. The proposed algorithms utilize color pyjamas for robust limb segmentation and tracking, unlike the purely motion-based method in [15]–[19].

3) Evaluation of the proposed system against EEG-based systems. The average length of studied video sequences is 109 s, where the longest sequence is 163 s. In contrast, the length in [12] is 14 s while that in [15]–[19] is only up to 20 s, and there is no evaluation against EEG-based systems in [12], [13], and [15]–[19].

II. Methods

A. Study Description and Experimental Setup

This study was conducted at the KK Women's and Children's Hospital, Singapore. We recruited those pediatric epileptic patients with informed consent obtained from their guardians. They were subjected to the standard video-EEG monitoring in the EMU. Moreover, we required them to wear color pyjamas and also mounted a second camera on the ceiling facing downward pointing to the center of patient bed for a full view. Considering user acceptability, we chose only two colors for the limb portions of pyjamas and this design was well accepted by our patients. The second camera records videos in color with 24-bit depth (RGB) at 12 frames/s with resolution of 384×288 pixels, saved in JPEG format for our study. In the standard EMU setting, light is only dimmed rather than completely turned OFF during night as the video-EEG system using an RGB camera as well. Furthermore, we have chosen bright colors (yellow and magenta) instead of darker colors for better capturing quality in dimmed light conditions. The auto-gain function of the second camera is also on for enhanced video quality. In addition, the EMU room temperature is kept at 20–24 °C (air-conditioned) such that blanket is not necessary at night.

Seizure events were identified from the standard video-EEG system by epileptologists. As a video-based approach, our system is limited to detect seizures with visible motor signs [3]. Thus, in this study, we consider only motor/hypermotor seizures. Five epileptic patients between the age of 1–15 years (four males and one female) were selected with 15 (motor) seizures in total. We obtained the time and duration of an interesting seizure event from the video-EEG system. We then took as a preseizure event (nonseizure immediately before the seizure event) of approximately equal duration. Hence, we have 15 sequences, where each sequence consists of a preseizure event followed by a seizure event. The average length of the 15 seizure sequences is about 109 s, which is significantly longer than 14 s in [12] and up to 20 s in [18] and [19]. The shortest sequence is about 34 s and the longest sequence is about 163 s.

Fig. 1 illustrates the proposed system. For each selected sequence, the user indicates a region of interest (the cyan rectangle in the figure) and the two foreground colors (the blue and green lines) using mouse, only for the first frame. In the figure, the patient's face is masked for anonymity and this is not part of the user initialization. After user initialization, the sequence is analyzed automatically. Limbs are detected first and then parameters are estimated to quantify their movements, forming movement trajectories, as shown in Fig. 1. We further analyze these trajectories to characterize the movements and detect seizures. These procedures will be described in more detail in the following sections.

B. Detection of Limbs From Videos

1) Foreground/Background Modeling: We employ the Gaussian mixture models (GMMs) to model foreground and background, as in [21]. A color video frame j of N pixels is represented by an array $\boldsymbol{\zeta}_j = (\zeta_1, \ldots, \zeta_n, \ldots, \zeta_N)$, where $\zeta_n = (R_n, G_n, B_n), n \in [1, N]$, which are the RGB values. Let an array $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N), \alpha_n \in \{0, 1\}$ denote the labeling of each pixel as background ($\alpha_n = 0$) or foreground ($\alpha_n = 1$). Two GMMs with K components are defined for background and foreground pixels, parameterized as [21]

$$\boldsymbol{\psi} = \{\pi(\alpha, k), \mu(\alpha, k), \Sigma(\alpha, k), \alpha = 0, 1, k = 1, \ldots, K\} \tag{1}$$

where π is the weight, μ is the mean, and Σ is the covariance matrix. The vector $\mathbf{k} = \{k_1, \ldots, k_n, \ldots, k_N\}$ with $k_n \in \{1, \ldots, K\}$ indicates the component of the background or foreground GMM that each pixel belongs to, assigning according to $\alpha_n = 0$ or 1. As the foreground is expected to be considerably simpler than the background, we keep $K = 6$ components for the foreground GMM, while $K = 12$ components for the background GMM.

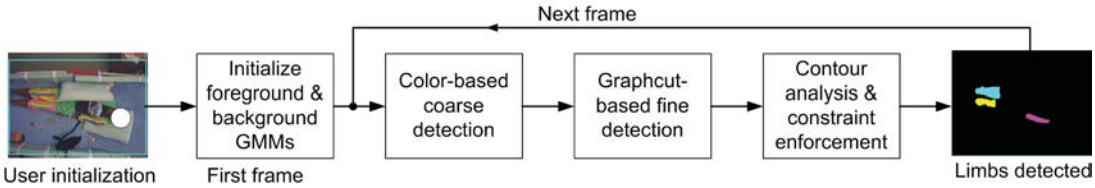

Fig. 2. Limbs are detected through color-based coarse detection and then Graphcut-based fine detection, followed by analysis of the contours and enforcing constraints. The patient's face is masked for anonymity.

2) Limb Detection: Our limb detection algorithm follows a coarse-to-fine approach adapted from [22] for efficiency, as shown in Fig. 2. For the first frame, the background and foreground GMMs are initialized based on user indicated foreground pixels. The coarse detection then assigns background/foreground labels to pixels based on colors using the following four rules:

1) Pixels with little color information (low saturation) are assigned as background ($\alpha_n = 0$).
2) Pixels with color very close (defined by the GMM likelihood [22]) to background colors (GMMs) in the first frame are assigned as background ($\alpha_n = 0$), for frame number $j > 1$.
3) Pixels with color very close to foreground colors (GMMs) in the first frame are assigned as foreground ($\alpha_n = 1$), for frame number $j > 1$.
4) Pixels corresponding to the erosions (by a 17×17 window) of limbs detected in the previous frame ($j - 1$) are assigned as foreground ($\alpha_n = 1$), for frame number $j > 1$.

The second and third rules are introduced to handle the cases where the limbs are not detected in a few previous frames, either occluded or out of view. They are based on the observation that with user initialization, the first frame segmentation is highly reliable and so are the corresponding GMMs. Rule 4 is an enforcement of the temporal constraint, under the assumption that a limb in the current frame should be close to the same limb in the previous frame.

After the coarse detection, the fine detection employs the popular GrabCut algorithm in [21] to produce foreground with large connected region of pixels with the same color. In GrabCut, a trimap Ω consists of three regions: Ω_B, Ω_F, and Ω_U, corresponding to the initial background, foreground, and uncertain pixels, respectively. Ω_B and Ω_F pixels are defined in the coarse detection, while pixels belonging to Ω_U are to be classified as either background or foreground. To classify pixels in Ω_U, $\boldsymbol{\alpha}$ is initialized as $\alpha_n = 0$ for $n \in \Omega_B$ and $\alpha_n = 1$ for $n \in \Omega_U \cup \Omega_F$. The background and foreground GMMs are initialized from $\boldsymbol{\alpha}$ with k-means clustering. Each pixel is then assigned a GMM component k_n and GMM parameters are learned. The segmentation is estimated through Graph Cut [20], [21], [23]. Finally, for the foreground GMMs, we run the k-means clustering algorithm on $\{\mu(1,k), k = 1, \ldots, 6\}$ to obtain two groups with different colors (as our pyjamas have two colors for the limbs). Foreground pixels belonging to each group are given the same label (1 or 2), resulting a segmentation map $\boldsymbol{\alpha}' = (\alpha'_1, \ldots, \alpha'_N)$, $\alpha'_n \in \{0,\ 1,\ 2\}$ denote the labeling of each pixel as background ($\alpha'_n = 0$), foreground 1 ($\alpha'_n = 1$), or foreground 2 ($\alpha'_n = 2$).

Next, we refine the detection by analyzing the segmented regions to obtain individual limbs using contour analysis techniques with the following constraints enforced:

1) For each limb, its area should be larger than a threshold (to remove noise and unreliable detections).
2) For each foreground color, there are at most two connected regions (corresponding to two limbs).
3) The position, orientation, and area of a limb detected in the current frame should not differ from those of the corresponding limb in the previous frame by a certain amount (to alleviate the correspondence problem).

In Rule 3 above, we assume that the movement between successive frames (i.e., within 0.08 s) is limited within a range. In addition, if a limb is not detected in the current frame, it is assumed to be fully occluded and the limb detected in the previous frame is taken as the current detection. Thus, a fully occluded limb is assumed to be still and it will not lead to any movement to trigger a detection. When a fully occluded limb reappears, the detection will need a few frames to recover. Unless all four limbs are fully occluded, limb movement information can be extracted from visible limbs for analysis. The initial segment labels are obtained based on the fixed design of pyjamas and known orientation of the patient body (with head on the right). The limbs in subsequent frames are labeled by finding the correspondence with limbs in the previous frames through examining interframe pairwise limb distances (defined by distance between their centroids).

Besides GrabCut [21], we also studied two other popular methods for detection and tracking: Histogram of Oriented Gradients (HOG) [24] and Scale-Invariant Feature Transform (SIFT) [25]. However, they did not give satisfactory results in our testing. The HOG-based method is designed for detecting objects with distinctive shape appearance, while in our context, we need to detect and localize individual limbs that have limited shape appearance distinctiveness. SIFT is a point-feature-based method that works well on images with rich texture, while our setting (a typical hospital ward) has many regions of homogenous colors lacking rich texture. Our GrabCut-based approach (combining GMM and Graph Cut) bears low computation cost and gives us the most satisfactory results in terms of efficiency and accuracy.

C. Movement Quantification for Seizure Detection

With limbs detected, we extract parameters to quantify their movements by adapting the approach in [26]–[28]. Dynamic parameters, i.e., the location (or position) and orientation (or angle) of a limb, are most relevant to limb movements [26]–[28].

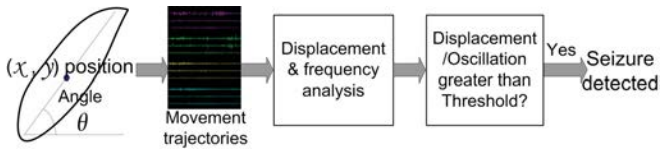

Fig. 3. Seizure can be detected by extracting discriminative features from limb movements described by three limb parameters for each limb: angle θ, x position, and y position.

Thus, a limb l is parameterized by three parameters: the angle θ (in degrees), the x position (in pixels), and the y position (in pixels), as shown in Fig. 3. For each frame j, there are 12 (3×4) movement parameters for four limbs:

$$\boldsymbol{\phi}_j = \{\theta(j,l), x(j,l), y(j,l), l = 1,2,3,4\}. \tag{2}$$

To estimate the three parameters of a limb, we fit an ellipse to the detected limb using the ellipse fitting method in OpenCV 2.3 [29], the x and y coordinates of the ellipse centroid are the x and y positions of the limb, respectively, and the orientation of the ellipse is the angle θ of the limb. When the fitted ellipse is close to a circle, the estimated orientation is not reliable and we set $\theta(j,l) = \theta(j-1,l)$ for $j > 1$. In addition, we enforce the following constraint for $j > 1$ to take temporal smoothness into account:

$$\begin{aligned} &|\theta(j,l) - \theta(j-1,l)| < \Gamma_\theta, \\ &|x(j,l) - x(j-1,l)| < \Gamma_x, |y(j,l) - y(j-1,l)| < \Gamma_y \end{aligned} \tag{3}$$

where we set the thresholds $\Gamma_\theta = \Gamma_x = \Gamma_y = 12$ in this study.

Next, we analyze limb movements to find distinct motion patterns for seizure detection. Based on our observations of limb trajectories for seizure and nonseizure events, we extract two features: the displacement feature in the time domain and the oscillation feature in the frequency domain. We perform displacement and frequency analysis on the trajectories. A seizure is considered detected if the displacement value (defined below) exceeds a threshold Γ_D or the oscillation strength (defined below) exceeds a threshold Γ_S. We analyze each trajectory using overlapping windows of 10 s (i.e., window length $W = 120$), with 1 s shift (i.e., 9 s overlap) so that we have a detection signal every second. Moreover, as the spatial resolution is limited at 384×288, our system is not able to detect subtle movements.

Displacement feature: For the displacement feature, we try to detect sustained displacement from a baseline in limb movements, which may relate to stiffening of limbs, e.g., in the tonic phase during seizures. We extract the displacement feature[1] from the differential values of each trajectory. The signal is preprocessed with a median filter of order 25 to remove noise. For a windowed differential signal $z(j), j = \beta, \ldots, \beta + W - 1$, the displacement value D is calculated as the accumulated differences below:

$$D = \left| \sum_{j=\beta+1}^{\beta+W-1} z'(j) \right| \tag{4}$$

[1]It should be noted that this limb displacement feature is a combination of torso movement and limb movement relative to the torso.

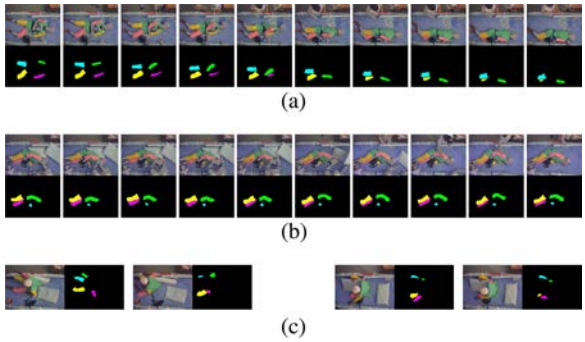

Fig. 4. Examples of detected limb movements. Note that all frames shown were captured during night with decent detection quality. The images are masked for anonymity. (a) During a seizure with sustained displacement around 4:45 A.M. (b) During a seizure with strong oscillation around 3:13 A.M. (c) Sitting positions around 3:12 A.M. and 3:56 A.M.

where

$$z'(j) = \begin{cases} z(j), & \text{if } |z(j)| \le 5 \\ sgn(z(j)) \cdot 5, & \text{otherwise.} \end{cases} \tag{5}$$

We impose a maximum change on $z(j)$ to get $z'(j)$ under the observation that sustained displacement due to stiffening should be gradual and smooth. A seizure is detected if $D > \Gamma_D$ in a window for at least three limb parameters.

Oscillation feature: For the oscillation feature, we aim to detect presence of strong oscillations in limb movements, which could be due to limb jerking, e.g., in the clonic phase during seizures. We extract the oscillation feature from the differential values of each trajectory. As we record videos at 12 frames/s, the sample frequency is 12 Hz. First, a differential trajectory is bandpassed to keep only the content between 1 and 5 Hz to remove noise. All movements within this band are considered with valid jerking frequency due to seizures, while movements beyond this band cannot be accurately detected. We then obtain the power spectrum through the fast Fourier transform. For each window, we examine the oscillation strength $S_{\max}$, which is defined in this paper as the magnitude of the frequency component that corresponds to the highest power of the spectrum. A seizure is detected if $S_{\max} > \Gamma_S$ in a window for at least one limb parameter.

III. Results and Discussions

A. Detection and Quantification of Limb Movements

With the proposed color-pyjamas-based video system, we can reliably detect body limbs of a patient for most of frames. Common detection problems are due to false detection of objects with colors similar to the limbs, or occlusions, including self-occlusion and occlusion by other subjects or objects. To illustrate, we show ten sample frames in the tonic phase of a seizure in Fig. 4(a), where sustained displacement is observed, particularly for the right arm. Similarly, ten sample frames in the clonic phase of a seizure are depicted in Fig. 4(b), where we can see limb oscillations due to jerking. We also show some results for sitting postures in Fig. 4(c). All frames in Fig. 4 are masked for anonymity.

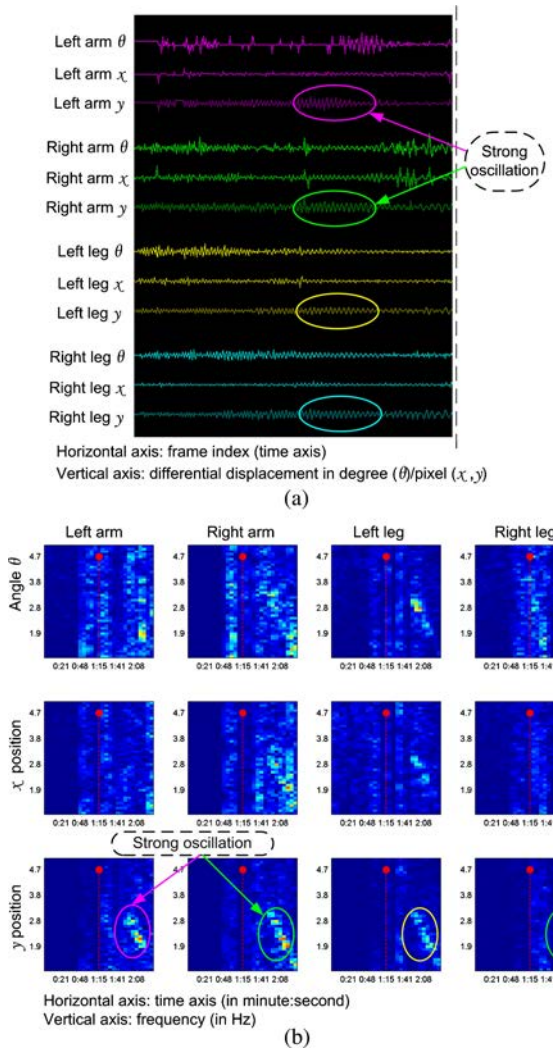

Fig. 5. Demonstration of the oscillation feature for the sequence partly shown in Fig. 4(b). Strong oscillation can be observed in both the spatial and frequency domains. (a) Strong oscillation in the movement trajectories. (b) Strong oscillation in the spectrograms.

To mimic EEG tracings for clinician's easy adaptation in reading, we follow [13] to display the quantified movement information as trajectories. Fig. 5(a) shows an example of the trajectories for 12 parameters from a short sequence. The horizontal axis is the time axis. The vertical axis indicates the differential values: $\theta(j,l) - \theta(j-1,l)$, $x(j,l) - x(j-1,l)$, and $y(j,l) - y(j-1,l)$. Thus, the trajectories visualize the changes in limb positions and orientations over time.

Next, we demonstrate the displacement and oscillation features for seizure detection in Figs. 6 and 5 for sequences partly shown in Fig. 4(a) and (b), respectively. In each subplot of 12 parameters in these figures, we use a red pole to indicate the ground-truth (manually labeled) of seizure onset time, which is around the center of the time axis. We highlight prominent features (sustained displacement and strong oscillation) with colored ellipses.

The presence of oscillation can be observed from high power bands in the spectrogram of affected limbs during seizures. We

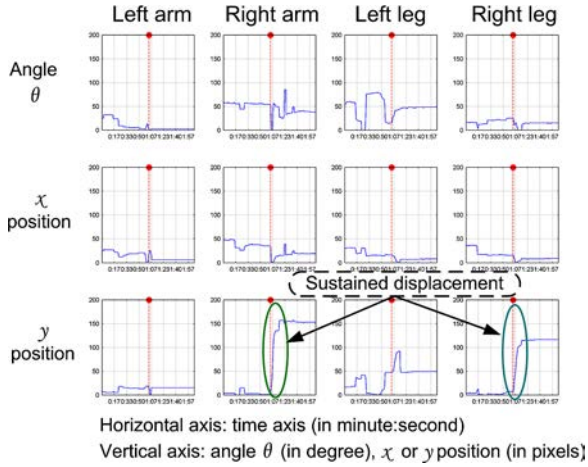

Fig. 6. Demonstration of the displacement feature for the sequence partly shown in Fig. 4(a). Sustained displacement can be captured by the movement parameters.

can see strong oscillations from several trajectories in Fig. 5(a) and also in the spectrograms for the y positions displayed in the third-row images of Fig. 5(b). In addition, the quantified limb movements reflect the frequency of jerking during epileptic seizures. For example, the jerking frequencies of y positions for four limbs decrease (slow down) from about 2.8 Hz to around 1.9 Hz in the interval highlighted by colored ellipses in Fig. 5(b), which cannot be accurately estimated through traditional visual inspection of video recordings or from EEG signals. To the best of our knowledge, this is the first demonstration of such capability for an unobtrusive video-based approach. This information could be useful in determining jerking pattern, severeness, or type of seizures.

Sustained displacement is most apparent on plots of raw parameters, as shown in Fig. 6. For example, the y position of the right arm (leg) has a displacement of about 150 (100) pixels upon seizure onset, as highlighted by colored ellipses in Fig. 6 [with sample frames shown in Fig. 4(a)], which cannot be accurately estimated through traditional visual inspection or from EEG signals. However, compared to oscillation, displacement feature is not as reliable since normal movements (e.g., normal turn/move in bed) can lead to sustained displacement as well so false positives are more likely.

B. Evaluation on Seizure Detection

We further study the feasibility of our color-based video system in seizure detection. We compare seizure detection results for the 15 sequences from five patients, as described in Section II-A. The start time, duration, and main posture(s) for each sequence are listed in Table I, where 9 out of 15 sequences were captured after sunset (7 P.M.) and before sunrise (7 A.M.). The limb detection parameters and the range of Γ_D and Γ_S are trimmed based on three seizures from patient B (B-1,B-2,B-3), and the same set of parameters are used for all patients for fair evaluation. Due to variations in age and physical size, the same parameters are not optimal for all patients.

TABLE I
START TIME, DURATION, AND MAIN POSTURE(S) OF THE 15 SEQUENCES

Index	Start time	Duration (sec)	Main postures
A-1	2:15PM	123	Lying
A-2	3:47PM	133	Lying
A-3	0:49AM	118	Lying
A-4	2:34AM	132	Lying
A-5	7:02PM	45	lying
A-6	11:18PM	34	Lying
B-1	3:11AM	163	Sitting,lying
B-2	3:56AM	157	Sitting,lying
B-3	4:44AM	125	Lying
C-1	2:07PM	144	Lying
C-2	0:02AM	146	Lying
D-1	9:33AM	86	Lying
E-1	6:57PM	70	Lying
E-2	6:13AM	97	Lying
E-3	12:50PM	70	Sitting,lying

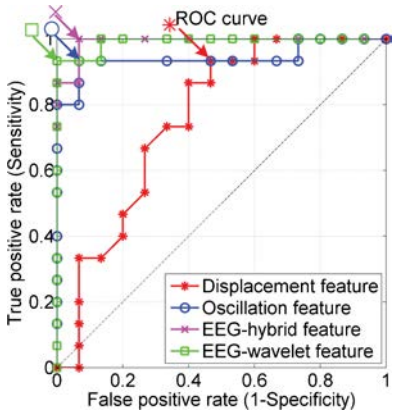

Fig. 7. ROC curves for the displacement feature, oscillation feature, EEG-hybrid feature, and EEG-wavelet feature. The displacement feature is the poorest, while the other three features have close performance.

We analyze two events (preseizure and seizure) for each sequence so there are 30 events (2×15 sequences) in total. When (at least) one seizure is detected in the seizure duration, we count it as a true detection, while when (at least) one seizure is detected in the preseizure duration, we count it as a false detection. The event labels provided by epileptologists serve as the ground truth.

We evaluate the detection performance of the proposed displacement and oscillation features extracted from videos using a simple threshold procedure. The performance varies for different values of Γ_D or Γ_S and we show the receiver operating characteristic (ROC) curves in Fig. 7. From the ROC curves, the displacement feature tends to have higher sensitivity, while the oscillation feature tends to have higher specificity.[2] This is expected as oscillations are unique to jerking and normal movements do not usually result in strong oscillation, while displacements often occur in nonseizure movements as well. In the comparison below, we chose operating points for these two features by maximizing the sum of their sensitivity and specificity, as indicated by the arrows with marker attached in Fig. 7, where $\Gamma_D = 35$ and $\Gamma_S = 2.2$.

[2]Sensitivity = true positive rate. Specificity = 1 − false positive rate.

TABLE II
SEIZURE DETECTION RESULTS FROM 15 SEQUENCES BY FOUR FEATURES: DISPLACEMENT, OSCILLATION, EEG-HYBRID [8], [30], AND EEG-WAVELET [4]

Patient	Displacement		Oscillation		EEG-Hybrid		EEG-Wavelet	
	TP	FP	TP	FP	TP	FP	TP	FP
A(6)	6	2	6	0	6	0	6	0
B(3)	3	2	2	0	3	0	3	0
C(2)	2	0	2	0	2	1	2	0
D(1)	0	0	1	0	1	0	1	0
E(3)	3	3	3	1	3	0	2	0
Total(15)	14	7	14	1	15	1	14	0

TP: true positive; FP: false positive. The displacement feature gives more false positives.

We benchmark our methods against two popular EEG-based methods: the hybrid seizure detection method (EEG-hybrid) in [8] and [30] and the wavelet-decomposition-based method (EEG-wavelet) in [4]. EEG signals in our EMU are recorded with 32 electrodes with standard international 10–20 system at 256-Hz sampling rate. The raw EEGs are preprocessed using double banana montage to result in 18 channels. For seizure detection, feature vectors are extracted from 10 s windows with 1 s shift (consistent with video-based methods). For the EEG-hybrid method, the montaged signals are bandpass-filtered to [0.5 Hz, 25 Hz]. Six univariate features (mean, entropy, curve length, energy, nonlinear energy, and spectral power) are summed up to form a power-based hybrid index for each channel. For the EEG-wavelet method, subband signals are obtained through seven-level decomposition using the Daubechies wavelet filter. Those from the fourth- to seventh-level decompositions represent activities at time-scales corresponding to the [0.5 Hz, 25 Hz] band that captures seizure onset of various electrographic manifestations [4]. Thus, the energy in each of these four subband signals is summed up to form a wavelet-energy-based index for each EEG channel. A seizure is detected if the power-based hybrid index or the wavelet-energy-based index is greater than a threshold Γ_h or Γ_w, respectively, in a window for at least ten channels. In performance comparison, the operating points for these two features are similarly chosen by maximizing the sum of their sensitivity and specificity, as indicated by the marked arrows in Fig. 7. From the ROC curves in Fig. 7, these EEG-based methods have better overall performance than the video-based methods, with the EEG-hybrid feature has higher sensitivity, while the EEG-wavelet feature has higher specificity.

We present the seizure detection results in Tables II–IV. Table II reports the number of true positives and false positives for each patient. The first column of Table II indicates the number of sequences/seizures for each patient in parentheses. The displacement feature is not unique to tonic seizures (stiffening) as many other significant nonseizure movements, such as normal turning in bed, can result in large displacement too. Thus, the displacement features have many false positives. On the other hand, the oscillation feature is characteristic of clonic seizures (jerking) as normal movements of a patient are less likely to give strong oscillation. Therefore, the oscillation feature gives only one false positive and one missed detection (out of 15 sequences). In contrast, the EEG-hybrid feature gives only one false detection, while the EEG-wavelet feature gives only one missed detection, showing better performance. Furthermore,

TABLE III
COMPARISON OF LATENCY IN SECONDS FOR TRUE SEIZURE DETECTIONS BY FOUR FEATURES: DISPLACEMENT, OSCILLATION, EEG-HYBRID [8], [30], AND EEG-WAVELET [4]

Seizure	A-1	A-2	A-3	A-4	A-5	A-6	B-1	B-2	B-3	C-1	C-2	D-1	E-1	E-2	E-3
Displacement	4	6	17	5	4	1	7	3	6	3	3	-	2	2	1
Oscillation	18	11	21	17	18	9	43	7	-	8	56	21	17	30	23
EEG-hybrid	25	12	13	29	0	4	67	97	73	20	0	22	15	16	24
EEG-wavelet	21	17	11	15	0	6	53	73	86	21	19	17	-	19	0

For missed detections, "-" is entered. The displacement feature has much lower latency than the others.

TABLE IV
COMPARISON OF SENSITIVITY (IN PERCENTAGE), SPECIFICITY (IN PERCENTAGE), AND AVERAGE LATENCY (IN SECONDS, FOR TRUE POSITIVES) BY FOUR FEATURES: DISPLACEMENT, OSCILLATION, EEG-HYBRID [8], [30], AND EEG-WAVELET [4]

Feature	Displacement	Oscillation	EEG-Hybrid	EEG-Wavelet
Sensitivity	93.3	93.3	100	93.3
Specificity	53.3	93.3	93.3	100
Latency	4.6	21.4	27.8	25.6

The displacement feature has poorer specificity but better latency.

the nonparametric McNemar's exact test [29], [30] has been performed to study the statistical significance of the detection results (30 events). The p values for the displacement feature paired with the oscillation, EEG-hybrid, and EEG-wavelet features are 0.0352, 0.0195, and 0.0195, respectively. The p values for the oscillation feature paired with the EEG-hybrid and EEG-wavelet features are both 0.500. This indicates that the performance difference (poorer) of the displacement feature is statistically significant from the oscillation, EEG-hybrid, and EEG-wavelet features at the significance level of 0.05. On the other hand, the performance difference of the oscillation feature is not statistically significant from the EEG-hybrid and EEG-wavelet features at the significance level of 0.05.

Table III reports the detection latency for each true seizure detection (named by [patient ID]-[seizure index]) in columns. "-" is entered for missed detections. The average latency together with sensitivity and specificity is summarized in Table IV. The displacement feature has low specificity, high sensitivity, and very low latency in detection. As mentioned earlier, this may be due to its less unique characteristics. Another cause is that in a motor seizure, the tonic phase usually comes first, followed by the clonic phase. Next, the oscillation feature has a lower latency than both EEG-based features, while its sensitivity and specificity are comparable to EEG-based features. Finally, the EEG-wavelet feature gives lower latency than the EEG-hybrid feature. On the whole, it is encouraging to see the performance of the oscillation feature to be close to the EEG-based features for the motor seizure sequences studied.

Hence, from their different characteristics, we may use both the displacement and oscillation features by giving *warnings* based on displacement features, while *alarms* based on the oscillation features. When caretakers are less occupied, they could react to warnings, while if caretakers are highly occupied, they could react to alarms only.

C. Discussions

On the basis of the results reported above, we have demonstrated the feasibility of using a color-based video system for epilepsy monitoring and seizure detection. The proposed method has the following three key benefits.

1) It can potentially serve as an additional tool in EMU for seizure detection when the patient is not continuously monitored by nurses, parents, or caretakers, especially during night. Moreover, in a clinical setting, the proposed video system can be combined with EEG-based detection/analysis. The two methods can be complementary. On one hand, when there are lots of movements, the video-based system has more clues, while EEG suffers more from artifacts. On the other hand, when there is little movement, the video-based system will fail, while EEG signals are usually clean to get correct detection.
2) Extraction of quantitative limb information on epileptic seizures provides a new method for investigating symptomatology during epileptic seizures, overcoming the disadvantages of visual inspection. For example, our system can estimate the jerking frequency and duration, but EEG or visual inspection cannot provide such quantitative measurements.
3) As only one camera needs to be mounted on the ceiling, the proposed system minimizes the introduction of cabling and clutter in the patient's living space, and it can be easily incorporated into personal computers [31], [32]. This makes it a promising solution for monitoring in a home environment.

On the other hand, as a new technology, we are still at the early stage of research on a video-based solution for seizure detection or patient monitoring in general. While the proposed system has many advantages, it is important to recognize its limitations as well as potential future works for respective enhancements when applicable:

1) At present, our system can work for seizures with large limb movements. It cannot be used to detect seizures with subtle movements or movements of eyes or head as the only signs. Furthermore, the present system will not work when the lighting condition is poor and colors cannot be differentiated reliably. To deal with low-light conditions, depth camera with emitted infrared light could be a promising solution. On the other hand, textures could be more reliable than colors in this case.
2) As we have been experimenting with different pyjamas designs in our investigation, the present system requires simple user initialization on the first frame. When the design is finalized, user initialization is not necessary when the background is clean (with respect to the limb colors) because of the customized color outfits. However, an initial checking of the video content is still beneficial to ensure

that the captured video is good for automatic analysis at the beginning.

3) Regarding the pyjamas design, various options can be explored. We can utilize four different colors for four different limbs to further lift the burden of segmentation and tracking of limbs, as well as mitigate the correspondence problem. In addition, smaller color patches (with possibly more colors) instead of a single color tube could give us more detailed information about limb movements, similar to that in [33]. However, we have to consider whether patients are willing to wear more colorful pyjamas. Our current designs are well accepted by pediatric patients tested so far.
4) In our present system, we model only 2-D movements of limbs using a single RGB camera, while limbs move in a 3-D space in fact. Therefore, movements along the camera optical axis cannot be detected. A further working direction is to develop 3-D models of limbs for more accurate quantification, e.g., utilizing stereo/multiple cameras. The Microsoft Kinect [34] has both RGB and depth cameras, which can complement each other. It also has a microphone array so it might be possible to further capture audio signals to detect abnormal sounds as well. Thus, Kinect-based system could be very promising.
5) Due to limitations of available suitable patients for analysis, we have tested only 15 motor seizure sequences in this study. By reporting our methods and interesting findings, we hope that those interested in video-based methods for clinical applications can benefit from our experience, thus accelerating research and development in this direction.

IV. Conclusion

We have described the implementation and evaluation of an unobtrusive video-based method for quantifying and characterizing limb movements in epilepsy monitoring. The proposed color-based video system uses a single ceiling camera with customized color pyjamas. After a simple user-initialization on the first frame, our system extracts the positions and angles of patient's limbs automatically. We further perform (time-domain) displacement analysis as well as frequency analysis to characterize limb movements and detect motor seizures. We identify sustained displacement and strong oscillation as two useful features. In experimental studies on 15 sequences from five patients, the oscillation feature has achieved performance comparable to EEG-based features, while the displacement feature is inferior. On the whole, the proposed video-based system is a promising approach for home monitoring and a good addition to investigation on quantified motor semiology for other movement disorders or behavioral changes, such as sleep disorder analysis.

Acknowledgment

The authors would like to thank the anonymous reviewers for their insightful and constructive comments.

References

[1] L. Iasemidis, "Epileptic seizure prediction and control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 5, pp. 549–558, May 2003.
[2] S. C. Schachter, J. Guttag, S. J. Schiff, D. L. Schomer, and S. Contributors, "Advances in the application of technology to epilepsy: The CIMIT/NIO epilepsy innovation summit," *Epilepsy Behavior*, vol. 16, no. 1, pp. 3–46, Sep. 2009.
[3] S. Noachtar and A. S. Peters, "Semiology of epileptic seizures: A critical review," *Epilepsy Behavior*, vol. 15, no. 1, pp. 2–9, Feb. 2009.
[4] A. Shoeb, H. Edwards, J. Connolly, B. Bourgeois, S. T. Treves, and J. Guttag, "Patient-specific seizure onset detection," *Epilepsy Behavior*, vol. 5, no. 4, pp. 483–498, 2004.
[5] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 41–56, Jan. 2008.
[6] H. Lu, H.-L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularized common spatial pattern with aggregation for EEG classification in small-sample setting," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 12, pp. 2936–2946, Dec. 2010.
[7] E. Acar, C. Aykukt-Bingol, H. Bingol, R. Bro, and B. Yener, "Seizure recognition on epilepsy feature tensor," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 4273–4276.
[8] A. B. Gardner, G. V. A. M. Krieger, and B. Litt, "One-class novelty detection for seizure analysis from intracranial EEG," *J. Mach. Learn. Res.*, vol. 7, pp. 1025–1044, 2006.
[9] E. Acar, C. Aykukt-Bingol, H. Bingol, R. Bro, and B. Yener, "Multiway analysis of epilepsy tensors," *Bioinformatics*, vol. 23, no. 13, pp. i10–i18, 2007.
[10] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 1, pp. 6–20, Oct. 2009.
[11] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A survey of multilinear subspace learning for tensor data," *Pattern Recognit.*, vol. 44, no. 7, pp. 1540–1551, Jul. 2011.
[12] L. Chen, X. Yang, Y. Liu, D. Zeng, Y. Tang, B. Yan, X. Lin, L. Liu, H. Xu, and D. Zhou, "Quantitative and trajectory analysis of movement trajectories in supplementary motor area seizures of frontal lobe epilepsy," *Epilepsy Behavior*, vol. 14, no. 2, pp. 344–353, Feb. 2009.
[13] Z. Li, A. M. da Silva, and J. P. S. Cunha, "Movement quantification in epileptic seizures: A new approach to video-EEG analysis," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 6, pp. 565–573, Jun. 2002.
[14] T. M. E. Nijsen, J. B. A. M. Arends, P. A. M. Griep, and P. J. M. Cluitmans, "The potential value of three-dimensional accelerometry for detection of motor seizures in severe epilepsy," *Epilepsy Behavior*, vol. 7, no. 1, pp. 74–84, Aug. 2005.
[15] N. B. Karayiannis, A. Sami, J. D. J. Frost, M. S. Wise, and E. M. Mizrahi, "Automated extraction of temporal motor activity signals from video recordings of neonatal seizures based on adaptive block matching," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 676–686, Apr. 2005.
[16] N. B. Karayiannis, B. Varughese, G. Tao, J. D. J. Frost, M. S. Wise, and E. M. Mizrahi, "Quantifying motion in video recordings of neonatal seizures by regularized optical flow methods," *IEEE Trans. Image Process.*, vol. 14, no. 7, pp. 890–903, Jul. 2005.
[17] N. B. Karayiannis, Y. Xiong, J. D. J. Frost, M. S. Wise, and E. M. Mizrahi, "Quantifying motion in video recordings of neonatal seizures by robust motion trackers based on block motion models," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 6, pp. 1065–1077, Jun. 2005.
[18] N. B. Karayiannis, G. Tao, J. D. F. Jr., M. S. Wise, R. A. Hrachovy, and E. M. Mizrahi, "Automated detection of videotaped neonatal seizures based on motion segmentation methods," *Clin. Neurophysiol.*, vol. 117, no. 7, pp. 1585–1594, 2006.
[19] N. B. Karayiannis, Y. Xiong, G. Tao, J. D. F. Jr., M. S. Wise, R. A. Hrachovy, and E. M. Mizrahi, "Automated detection of videotaped neonatal seizures of epileptic origin," *Epilepsia*, vol. 47, no. 6, pp. 966–980, 2006.
[20] H. Lu, H.-L. Eng, B. Mandal, D. W. S. Chan, and Y.-L. Ng, "Markerless video analysis for movement quantification in pediatric epilepsy monitoring," in *Proc. 33th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug./Sep. 2011, pp. 8275–8278.
[21] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, Aug. 2004.
[22] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Coarse-to-fine pedestrian localization and silhouette extraction for the gait challenge data sets," in *Proc. Int. Conf. Multimed. Expo.*, Jul. 2006, pp. 1009–1012.
[23] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.

[24] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, vol. 1, pp. 886–893.
[25] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[26] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A full-body layered deformable model for automatic model-based gait recognition," *EURASIP J. Adv. Signal Process.*, vol. 2008, pp. 1–14, 2008.
[27] L. Lee, G. Dalley, and K. Tieu, "Learning pedestrian models for silhouette refinement," in *Proc. Int. Conf. Comput. Vis.*, Oct. 2003, pp. 663–670.
[28] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "A layered deformable model for gait analysis," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recognit.*, Apr. 2006, pp. 249–254.
[29] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision With the OpenCV Library*. Cambridge, MA: O'Reilly Media, 2008.
[30] Y. Pan, S. S. Ge, F. R. Tang, and A. A. Mamun, "Detection of epileptic spike-wave discharges using SVM," in *Proc. 2007 IEEE Int. Conf. Control Appl.*, Oct. 2007, pp. 467–472.
[31] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.
[32] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Exp.*, vol. 18, no. 10, pp. 10762–10774, 2010.
[33] R. Y. Wang and J. Popović, "Real-time hand-tracking with a color glove," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 1–8, Aug. 2009.
[34] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2011, pp. 1297–1304.

Haiping Lu (S'02–M'09) received the B.Eng. and M.Eng degrees in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2001 and 2004, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2008.

He is currently a Scientist at the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore. He is a coauthor of the book *Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data* (New York: CRC Press, 2013). His current research interests include machine learning, pattern recognition, and biomedical engineering.

Dr. Lu received the 2013 IEEE Computational Intelligence Society Outstanding Ph.D. Dissertation Award.

Yaozhang Pan (S'05–M'09) received the Ph.D. degree from the National University of Singapore, Singapore, in 2010, and the B.Eng. and M.Eng. degrees from the Harbin Institute of Technology, Harbin, China, in 2004 and 2006, respectively.

She is currently a Scientist at the Institute for Infocomm Research, Agency for Science, Technology and Research. Her research interests include machine learning, pattern recognition, EEG signal processing for medical applications, and brain–computer interface.

Bappaditya Mandal received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT), Roorkee, India, and the Ph.D. degree in electrical and electronic engineering from Nanyang Technological University (NTU), Singapore, in 2003 and 2008, respectively.

He is currently a Scientist at the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore. His research interests include subspace learning, feature extraction and evaluation, computer vision, image and signal analysis, and machine learning.

Dr. Mandal was a recipient of the Best Biometric Student Paper Award at the 19th International Conference on Pattern Recognition, 2008 and a recipient of full Research Scholarship Award from NTU, between 2004 and 2008. In 2001, he received the Summer Undergraduate Research Award from IIT.

How-Lung Eng (M'03) received the B.Eng. and Ph.D. degrees in electrical and electronic engineering from Nanyang Technological University, Singapore, in 1998 and 2002, respectively.

He is currently a Research Scientist with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore. His research interest includes real-time vision, pattern classification and machine learning for abnormal event detection. He has made several PCT fillings related to video surveillance applications and has actively published his works in the above areas of interest.

Dr. Eng was a recipient of the Tan Kah Kee Young Inventors' Award 2000 (Silver, Open Section) for his Ph.D. work, and a recipient of the TEC Innovator Award 2002, the IES Prestigious Engineering Awards 2006 and 2008, and IWA Project Innovation Award 2012 for his works in the areas of visual surveillance.

Cuntai Guan (S'91–M'92–SM'03) received the Ph.D. degree in electrical and electronic engineering from Southeast University, Nanjing, China, in 1993.

He is currently a Principal Scientist at the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore. He is also the Department Head of Neural & Biomedical Technology Department at Institute for Infocomm Research. He has contributed to more than 150 refereed journal and conference papers and holds 13 granted patents and applications. His research interests include neural and biomedical signal processing; neural and cognitive process and its clinical application; and brain–computer interface algorithms, systems, and its applications.

Dr. Guan is an Associate Editor for the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, and Associate Editor for *Frontiers in Neuroprosthetics*.

Derrick W. S. Chan received the Medical degree from Nottingham University, Nottingham, U.K., and the Master's degree in clinical investigation from the National University of Singapore, Singapore. He was trained in Paediatric EEG and Epilepsy in Melbourne, Vic., Australia and Toronto, ON, Canada, respectively.

He is currently the Head and Consultant in Paediatric Neurology at KK Women's and Children's Hospital, Singapore. His research interests include the epidemiology and pharmacotherapy of epilepsy, neurophysiology, and applications of advanced technology in clinical practice for paediatric epilepsy.