Extraction of Motion Strength and Motor Activity Signals From Video Recordings of Neonatal Seizures

Nicolaos B. Karayiannis, *Member, IEEE*, Seshadri Srinivasan, Rishi Bhattacharya, *Student Member, IEEE*, Merrill S. Wise, James D. Frost, Jr., and Eli M. Mizrahi

***Abstract*—This paper presents two methods developed to extract quantitative information from video recordings of neonatal seizures in the form of temporal motion strength and motor activity signals. Motion strength signals are extracted by measuring the area of the body parts that move during the seizure and the relative speed of motion using a combination of spatiotemporal subband decomposition of video, nonlinear filtering, and segmentation. Motor activity signals are extracted by tracking selected anatomical sites during the seizure using a modified version of a feature-tracking procedure developed for video, known as the Kanade–Lucas–Tomasi (KLT) algorithm. The experiments indicate that the temporal signals produced by the proposed methods provide the basis for differentiating myoclonic from focal clonic seizures and distinguishing these types of neonatal seizures from normal infant behaviors.**

***Index Terms*—Focal clonic seizure, motion strength signal, motor activity signal, myoclonic seizure, neonatal seizure, spatiotemporal video decomposition, temporal feature tracking, video processing and analysis.**

I. Introduction

NEONATAL seizures are often the first and, in some situations, the only clinical sign of central nervous system dysfunction in the newborn. Identification of seizures in the newborn initiates a prompt evaluation for a wide range of etiologies and, whenever possible, treatment of the underlying pathological processes. In some situations, antiepileptic medication is provided to diminish the likelihood of recurrent seizures and to lower the risk of physiologic instability during seizures. The presence of seizures may also affect prognosis, particularly with regard to neurodevelopmental sequelae and risk for certain forms of epilepsy. Thus, prompt recognition of seizures by nursery personnel is important with regard to diagnosis and management of underlying neurological problems.

Despite the importance of seizure recognition, most neonatal intensive care units and nurseries have limited resources for seizure identification. The attention of nursing personnel is distributed across a large number of newborns, many of whom are ill and require continuous bedside care. Neonatal seizures are often brief and may not be recognized since nurses and physicians cannot provide continuous surveillance of all infants in the neonatal intensive care unit. In addition, although neonatal intensive care unit nurses are highly trained in many aspects of care, there is significant variability in the level of skill and experience in seizure recognition among nurses. These factors illustrate the clear need for improved seizure surveillance methods that supplement direct observation by nurses and physicians and that are practical and economically feasible for routine use in the neonatal intensive care environment.

Early attempts to characterize neonatal seizures involved primarily bedside observation and relatively brief electroencephalogram (EEG) recordings. The more recent development of portable EEG/video/polygraphic monitoring techniques allows investigators to assess and characterize neonatal seizures at the bedside and permits retrospective review [17]. Investigations using these techniques have confirmed that the majority of neonatal seizures are either electroclinical (electrographic and clinical features that are temporally linked) or clinical only (clinical features with no consistent electrographic correlate) in character [27], [29], [49]. These techniques are generally used for only a few hours of monitoring and are not routinely available in many centers. Most research involving neonatal seizures has focused on analysis of EEG features and no investigations have used quantitative techniques to characterize visual features. This observation contrasts with the fact that the majority of seizures in the newborn are clinically expressed, either with or without an electrographic signature. Thus, automated video processing and analysis may supplement and extend human analysis of clinical seizure behaviors and may provide new information leading to more useful classification schemes.

Video recording is typically used with synchronized EEG and other polygraphic measures to analyze the characteristics of a clinical seizure after its recording [1], [3]–[5], [8], [10], [11], [17]–[19], [22], [24], [28], [31]–[33], [35], [47]. This technique is limited in terms of duration of recording and the availability of trained physicians for analysis. However, post-seizure analysis in the neonate can facilitate the classification of the event as epileptic or nonepileptic, determine the type of the ictal event (e.g., clonic, tonic, myoclonic, motor automatisms, and spasms), determine the EEG localization and associated clinical features of onset and evolution (focal, generalized, multifocal, alternating, migrating, etc.), reveal the precise sequence of motor components within a single seizure, and establish the temporal relationship of the observed motor activity to EEG activity.

Manuscript received September 5, 2000; revised June 26, 2001. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was M. W. Vannier. *Asterisk indicates corresponding author.*

*N. B. Karayiannis is with the Department of Electrical Computer Engineering, University of Houston, N308 Engineering Building 1, Houston, TX 77204-4005 USA (e-mail: karayiannis@uh.edu).

S. Srinivasan and R. Bhattacharya are with the Department of Electrical Computer Engineering, University of Houston, Houston, TX 77204-4005 USA.

M. S. Wise, J. D. Frost, Jr., and E. M. Mizrahi are with the Department of Neurology, Baylor College of Medicine, Houston, TX 77030 USA.

Publisher Item Identifier S 0278-0062(01)08667-0.

Most studies using video as a diagnostic tool for seizure characterization deal with problems associated with video recording itself. Synchronization has been one of the major practical obstacles for the simultaneous recording of EEG and motor activity [11]. The synchronization problem has been overcome by the development of integrated recording systems capable of recording motor activity as depicted on video, along with synchronized EEG and projecting both on a split screen.

This paper presents video processing and analysis techniques developed to extract quantitative information regarding the behavioral characteristics of neonatal seizures. More specifically, this paper describes two methods developed to extract temporal motion strength and motor activity signals from video recordings of neonatal seizures using a combination of spatiotemporal subband decomposition of video and two-dimensional (2-D) tracking of selected anatomical sites.

II. Video Processing and Analysis of Neonatal Seizures

The extraction of temporal motion strength and motor activity signals from video recordings of neonatal seizures is the first step toward the development of an automated video processing and analysis system for use in clinical settings [39]. A video system based upon automated analysis potentially offers a number of advantages. Infants in the neonatal intensive care unit could be monitored continuously using relatively inexpensive and noninvasive video techniques that would supplement direct observation by nursery personnel. This would represent a major advance in seizure surveillance and offers the possibility for earlier identification of potential neurological problems and subsequent intervention. From a research perspective, automated video processing and analysis holds great potential for refined characterization of clinical events. Characterization of clinical events has relied primarily upon visual analysis and consensus among pediatric neurologists, neonatologists, and clinical neurophysiologists regarding which paroxysmal behaviors represent clinical seizures. This has contributed to controversy regarding definitions of neonatal seizures and at times even skilled and experienced clinical neurophysiologists have different opinions regarding whether a specific behavior represents seizure activity.

Quantitative analysis using computerized video techniques may supplement and extend human analysis and may generate novel methods for extracting relevant information from paroxysmal behaviors. In certain types of neonatal seizure behaviors, refined analysis may shed light on specific motor activity patterns or attributes that constitute seizures, as compared with repetitive behaviors that do not represent seizures and do not have the same clinical relevance. Development of a quantitative, computerized method could lead to a more rigorous definition of neonatal seizures and could uncover key motor signatures that are not recognized using traditional visual analysis or limited monitoring of body/limb motion by EMG or accelerometry. Specific examples include: 1) the differentiation of focal clonic seizures from other repetitive movements such as tremor or semirhythmic nonpurposeful movements and 2) the assessment of movement characteristics in myoclonic seizures, such as amplitude and velocity of movements and synchrony of movements between left and right extremities. These examples represent common occurrences in the clinical setting and automated video techniques directed at analysis of specific components of movement will contribute to a more objective and quantitative analysis. In addition, these techniques may provide the basis for further work directed toward understanding the pathophysiology of certain seizure behaviors (epileptic versus nonepileptic mechanisms) and formulating more refined capabilities regarding prediction of outcome based upon the clinical presentation of neonatal seizures.

Recent developments in video processing and analysis research can facilitate the analysis of neonatal seizures. These developments have been stimulated by the transition from analog to digital video, which is expected to expand the use of computing devices into video processing and analysis. One of the major problems associated with video processing and analysis is the huge amount of data involved. Nevertheless, video recordings of neonatal seizures are highly redundant since infants may not move excessively in their beds while focal clonic and myoclonic seizures affect specific parts of their bodies, such as their extremities. Thus, the extraction of quantitative information from video-taped seizures must focus only on the moving parts of the infant's body that are relevant to the seizure. Quantitative analysis of video-taped neonatal seizures requires the development of mechanisms for tracking and quantifying motion of the infant's body parts during the seizure. This can be accomplished by two different, but complementary, methods proposed in this paper.

III. Extraction of Temporal Motion Strength Signals

The extraction from video recordings of quantitative information that is relevant only to the seizure can be accomplished by exploiting the redundancy typically found in video signals, namely the redundancy between adjacent frames (interframe redundancy) and the redundancy within each frame (intraframe redundancy). The intraframe and interframe redundancy can be utilized for identifying the infant's moving parts by performing spatiotemporal subband decomposition of the image sequences that compose the video recording. Subband decomposition allows the processing and analysis of signals, images, and image sequences (i.e., video) at different resolutions from a set of frequency selective subbands [7], [13], [15], [16], [23], [34], [43], [46]. The spatiotemporal decomposition of an image sequence begins with temporal decomposition, which is followed by spatial decomposition of the resulting temporal subbands. Temporal decomposition of image sequences is typically performed by a filter of length two, i.e., the shortest possible nontrivial filter, in order to minimize the computational burden associated with temporal filtering. Spatial decomposition is typically performed by longer wavelet filters in order to improve the frequency selectivity of the resulting subbands.

Fig. 1 illustrates an 11-band spatiotemporal subband decomposition of a video signal. In the temporal decomposition phase, the frames of the image sequence are passed block-by-block through a filter bank containing a low-pass temporal (LP_t) filter and a high-pass temporal (HP_t) filter. If temporal decomposition is performed by a filter of length two, each block

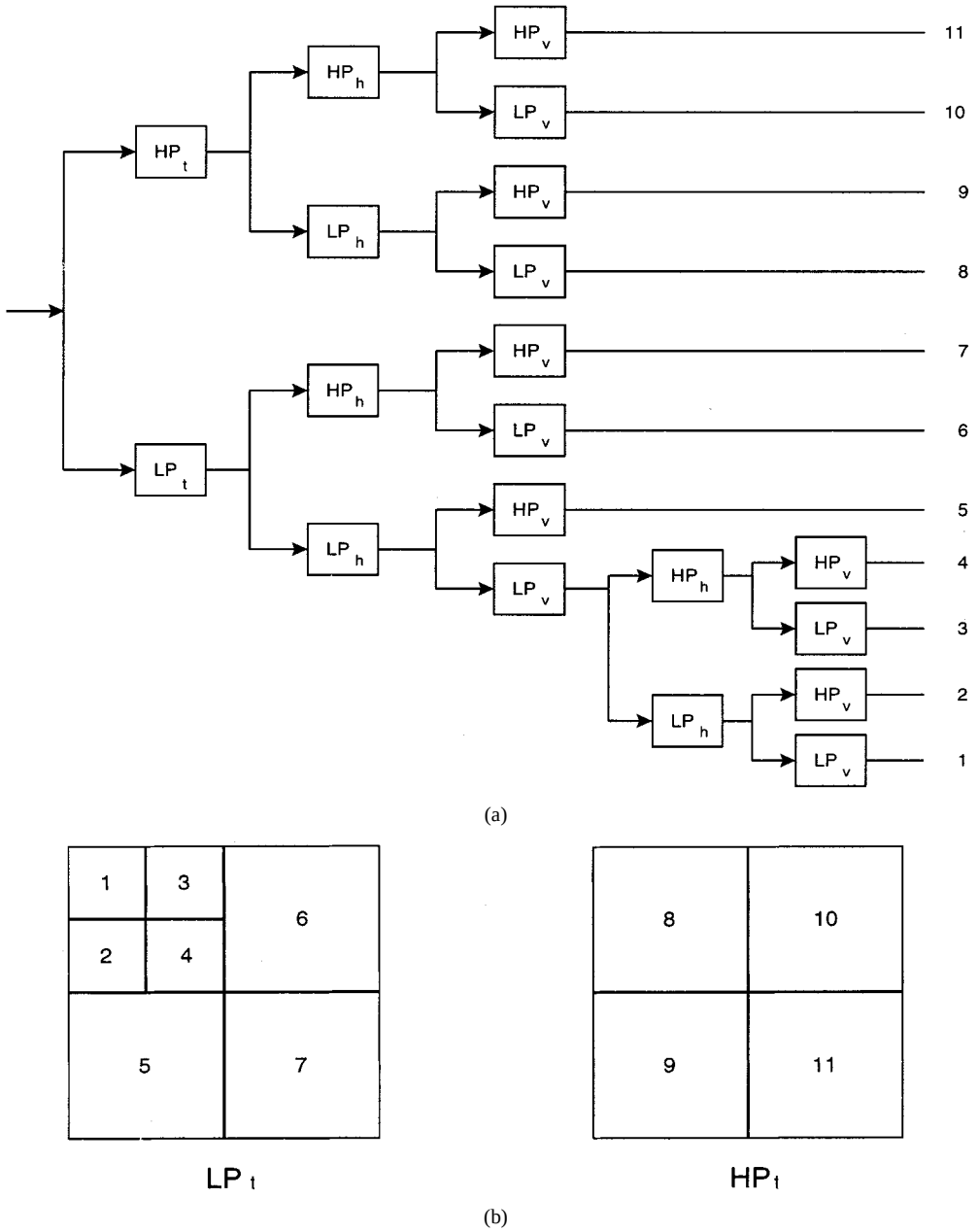

Fig. 1. Spatiotemporal subband decomposition of image sequences based on wavelets.

contains two consecutive frames of the sequence. Therefore, the temporal filtering results in two subbands: the low-pass temporal (LPT) subband and the high-pass temporal (HPT) subband. The LPT subband is computed by averaging two successive frames of the image sequence and carries low frequencies. The HPT subband represents the difference between two successive frames and thus, can be used to detect motion. In the spatial decomposition phase, each of the LPT and HPT subbands is passed through a filter bank which performs low-pass filtering along the horizontal dimension ($\mathrm{LP_h}$) and high-pass filtering along the horizontal dimension ($\mathrm{HP_h}$). Low-pass and high-pass filtering is followed by downsampling by a factor of two. Each of the resulting subbands is passed through a filter bank which performs low-pass filtering along the vertical dimension ($\mathrm{LP_v}$) and high-pass filtering along the vertical dimension ($\mathrm{HP_v}$). Low-pass and high-pass filtering is once again followed by downsampling by a factor of two. This sequence of operations completes one level of spatial decomposition of the LPT and HPT subbands. If necessary, the resulting subbands can be further decomposed. As an example, Fig. 1(a) describes an additional level of spatial decomposition of the subband produced by the $\mathrm{LP_t}$, $\mathrm{LP_h}$, and $\mathrm{LP_v}$ filtering operators. Fig. 1(b) shows the 11 subbands produced by the spatiotemporal decomposition scheme described in Fig. 1(a). The LPT subband produces seven subbands after two levels of spatial decomposition. Subbands 5, 6, and 7 result directly from the first level of spatial decomposition. Subbands 1–4 are the result of the second decomposition level applied on the upper-left subband, which is produced by the first decomposition level and contains low frequencies in both horizontal and vertical dimensions. Subbands 8–11 are produced by applying one level of spatial decomposition on the HPT subband. The

time-sequence composed of subband 1 is a low-resolution (both spatially and temporally) version of the original video. Because of its significant role, subband 1 is frequently called the dominant subband. The time sequences composed of subbands 2–11 are auxiliary video signals containing high-frequency detailed information. The nondominant subbands 2–11 are sparse and highly structured. Subband 8 contains the low-frequency components of the HPT subband in both horizontal and vertical dimensions. As a result, this particular subband carries most of the energy among subbands 8–11 produced by the decomposition of the HPT subband [43]. Subband 8 is often used as a motion detector instead of the HPT subband since it contains most of the information carried by the HPT subband and is reduced in size by a factor of 1/4 [13].

Temporal motion strength signals were extracted from video-taped neonatal seizures by measuring the area of the body parts moving between successive frames and the relative speed of motion. This method relied on the spatiotemporal decomposition of the image sequence that constitutes the video recording. Temporal decomposition was performed by the Haar filter of length two while spatial decomposition was performed by the Daubechies wavelet filter of length 20 [7]. Motion was detected and measured on subband 8 of the decomposed image sequence, which detects motion between successive frames of the sequence. Fig. 2(b) shows subband 8 computed on the four frames of the video-taped myoclonic seizure shown in Fig. 2(a). Subband 8 corresponding to frame 14 shows clearly the infant's left leg, which moves to the right and toward the top of the frame between frames 10 and 16 (Fig. 2 shows only frame 14). The infant's left leg is not visible in subband 8 computed on frames 100 and 200 since there was no motion between frames 99 and 100 or between frames 199 and 200.

The experiments indicated that subband 8 contains the moving body parts, but it is also corrupted by spiky noise, probably due to camera jitter and other recording imperfections. The noise appears as spurious patches (i.e., spikes) that occupy very small areas in comparison with those of the moving body parts. Most of these spurious patches were removed from subband 8 by applying median filtering [12], a computationally simple nonlinear operator that is particularly effective for this kind of noise. More specifically, subband 8 was filtered using a 2-D median filter of size 3×3 pixels. It was found that this filter size guarantees sufficient noise removal without any noticeable blurring effect on the moving body parts. Fig. 2(c) shows the frames shown in Fig. 2(b) after median filtering. Median filtering eliminated most of the spurious patches appearing in Fig. 2(b). As a result, the infant's left leg is clearly traced in frame 14. However, frame 100 contains some spurious clusters of pixels even after median filtering.

Following median filtering, the time sequence constituted by subband 8 was segmented in order to isolate the moving body parts from background noise and other spurious clusters of pixels that have typically lower intensity values than those of the moving body parts. Segmentation was performed by an adaptive version of the c-means (or k-means) algorithm [9]. This adaptive clustering algorithm clustered all pixels of each frame from the sequence formed by subband 8 in $c = 3$ clusters. The clusters produced by the c-means algorithm on the previous frame were used to initialize the clustering procedure for the current frame. This was done in order to speed up the segmentation process by exploiting the correlation between adjacent frames. Following the clustering process, one of the three clusters produced for each frame contains the pixels belonging to moving body parts, while the other two clusters contain background pixels as well as pixels representing background and other spurious information. The segmentation process was completed by assigning to all pixels belonging to the cluster of the highest intensities the same intensity value of 255 (corresponding to white color in a black-and-white image). All other pixels were assigned the intensity value of zero (corresponding to black color in a black-and-white image). Thus, the segmentation process produced a sequence of black-and-white frames that display the moving body parts as white areas in a black background. Fig. 2(d) shows the four frames produced by segmenting the frames shown in Fig. 2(c). Segmentation eliminated all spurious clusters of pixels in frames 0, 100, and 200, which contained no moving body parts. Segmentation also eliminated the low-intensity clusters of pixels from frame 14, which led to a better definition of the moving body part. The traces of the infant's left leg are shown in frame 14 as white patches in a black background.

The experimental results indicated that the segmented frames may still contain a few spurious bright patches due to noise in the original video recording. The contribution of such spurious patches to the measurements extracted from video recordings was prevented in this study by tracking the centroids of the bright patches in the frame sequence produced by the segmentation process. Tracking was performed by considering only those areas whose centroids were present within a small radius between successive frames. Averaging all such areas over successive frames produced the temporal signal $A_{\mathrm{av}}(t)$, which measures the average area occupied by the moving body parts over time. The experiments also indicated that seizure identification and recognition may benefit by a scaling scheme that can magnify fast motion of small body parts while suppressing slow motion of bigger body parts that may not be caused by a seizure. Scaling was performed in this study by multiplying the areas of the moving parts by the distance covered by them between adjacent frames. This scheme produced the temporal signal $A_{\mathrm{sc}}(t)$, which depends rather heavily on motion speed. The signal $A_{\mathrm{sc}}(t)$ can potentially facilitate the identification of seizures involving small body parts.

IV. Extraction of Temporal Motor Activity Signals

Measuring the motor activity of certain parts of the human body is a challenging problem due to the amount of data produced by locomotion in a three-dimensional (3-D) space. This problem is often simplified by data reduction techniques based on the projection of selected kinematic data into a space of lower dimensionality [6], [30], [40]–[42], [48], [50]. Neonatal seizures occur in a 3-D space, but infants viewed in bed by a video system are confined in a 2-D plane. In this application, data reduction can be accomplished by projecting the location of selected anatomical sites to the horizontal and vertical axes. As the seizure progresses in time, these projections will produce

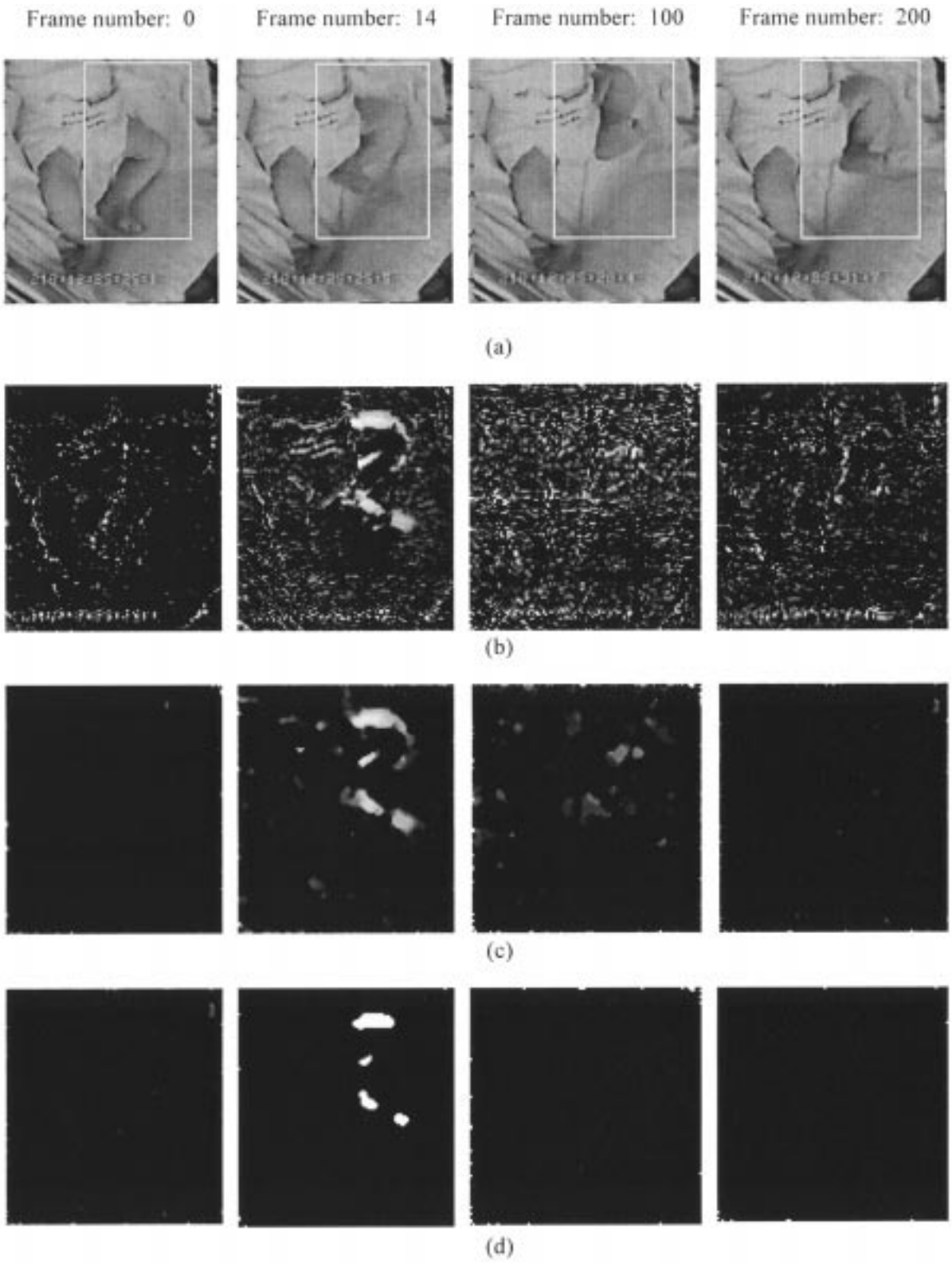

Fig. 2. Extraction of temporal motion strength signals. (a) Selected frames from a video-taped myoclonic seizure. (b) Frames produced by computing subband 8 using spatiotemporal decomposition. (c) Frames produced by applying median filtering on subband 8. (d) Frames produced by segmenting the filtered version of subband 8.

temporal signals recording motor activity of the body parts of interest.

Fig. 3 illustrates the mechanism that was used for extracting temporal signals tracking the movements of different parts of the infant's body during focal clonic and myoclonic seizures. Fig. 3 depicts a single frame containing the sketch of an infant's body with four selected anatomical sites. In this particular configuration, X_{LL} and Y_{LL} represent the projections of the site located

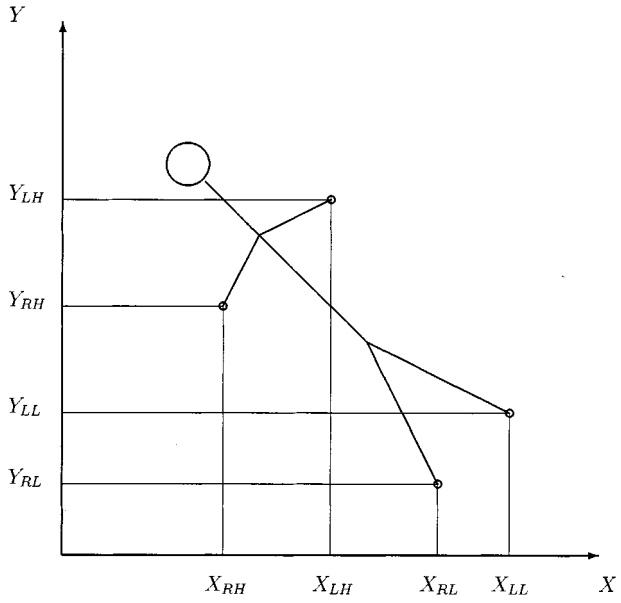

Fig. 3. Extration of temporal motor activity signals by projecting four sites of an infant's body to the horizontal and vertical axes.

at the left leg to the horizontal and vertical axes, respectively. The projections of the sites located at the right leg, left hand, and right hand are denoted by X_{RL} and Y_{RL}, X_{LH} and Y_{LH}, X_{RH} and Y_{RH}, respectively. As the infant moves its extremities, the locations of the sites in the frame will change, as will the projections of the sites to the horizontal and vertical axes. Recording the values of the projections from frame to frame of the video-taped seizure will generate four pairs of temporal signals, namely the signals $X_{LL}(t)$ and $Y_{LL}(t)$ for the left leg, the signals $X_{RL}(t)$ and $Y_{RL}(t)$ for the right leg, the signals $X_{LH}(t)$ and $Y_{LH}(t)$ for the left hand, and the signals $X_{RH}(t)$ and $Y_{RH}(t)$ for the right hand. For a given set of anatomical sites, each seizure will produce signature signals depending on its type and location.

Temporal motor activity signals were extracted from video-taped neonatal seizures by projecting a selected anatomical site to both horizontal and vertical axes. This method relied on an automated algorithm developed to track the site of interest in successive frames of video-taped neonatal seizures. The site-tracking algorithm was developed in this study by modifying and extending the KLT algorithm, a feature-tracking procedure developed for video by Tomasi and Kanade [44] based on earlier work by Lucas and Kanade [21]. The KLT algorithm automatically selects "good features" from the first frame of an image sequence. A good feature is one that can be tracked well throughout the entire image sequence. The selection of good features is based on the requirement that the spatial gradient matrix computed on the corresponding frame location is above noise level and well conditioned. The noise requirement implies that both eigenvalues of the gradient matrix must be sufficiently large, while the conditioning requirement means that the eigenvalues cannot differ by several orders of magnitude. Thus, a window is accepted as a good feature if the two eigenvalues λ_1 and λ_2 of the corresponding gradient matrix satisfy the condition $\min\{\lambda_1, \lambda_2\} > \lambda$, where λ is a predetermined threshold. The features selected from the first frame are then tracked through the image sequence by using a Newton–Raphson optimization method to minimize the difference between the windows in successive frames. The tracking scheme was improved by Shi and Tomasi [38], who extended the Newton–Raphson search method to operate under affine image transformations. This modification led to an optimal feature selection criterion and a feature monitoring scheme that can detect occlusions, disocclusions, and points that do not correspond to visually important features.

The latest version of the KLT algorithm was utilized in this study to track selected anatomical sites in video-taped neonatal seizures. Fig. 4(b) shows the location of 400 features selected and tracked by the KLT algorithm in the four frames of the video-taped myoclonic seizure shown in Fig. 4(a). It is clear from Fig. 4(b) that the features selected by the KLT algorithm in the first frame of the sequence (i.e., frame 0) are almost uniformly distributed over the entire frame area. However, the KLT algorithm became increasingly selective as the seizure progressed. In frames 14, 100, and 200, the features are located at the frame area occupied by the infant's body (including the moving body part) and the textured and nonhomogeneous areas of the background. It is also remarkable that the KLT algorithm did not track throughout the entire sequence the features located between the infant's legs at the lower-right quadrant of the frames. This can be attributed to the uniform intensity profiles of the windows considered by the algorithm in this homogeneous area.

Although the KLT algorithm was generally successful, in some cases the algorithm lost some features that were particularly important for this application. Lost features are typically tracked by the algorithm in the initial sequence of frames, but are not selected for tracking in subsequent frames. The experiments also indicated that the KLT algorithm failed to track moving body parts through the entire frame sequence when those parts contained a large amount of lost features. The susceptibility of the KLT algorithm to "lost features" motivated the tracking of a sufficiently large number of features within a predetermined radius from the selected anatomical site in the frame sequence. Fig. 4(c) shows the features tracked by the KLT algorithm within a predetermined radius from the site in the infant's left foot. In the first frame of the sequence (i.e., frame 0), the KLT algorithm selected four features within the neighborhood of the site indicated by a circle. It is apparent from Fig. 4(c) that one of the original features has been lost in frame 14. Only two of the features in the neighborhood of the site were tracked by the KLT algorithm in subsequent frames of the sequence. One of these features provided the reference for tracking the site through the entire sequence.

The strategy described above allowed the tracking of the site through the frame sequence even in cases where some of the features in its close neighborhood were lost by the KLT algorithm at some point in time. When the radius was sufficiently small, there were no noticeable differences between the ideal temporal signals (i.e., the signals corresponding to projections of the site) and the resulting temporal signals (i.e., the signals corresponding to the projections of the features tracked by the algorithm). This is clear from Fig. 4(d), which shows the feature

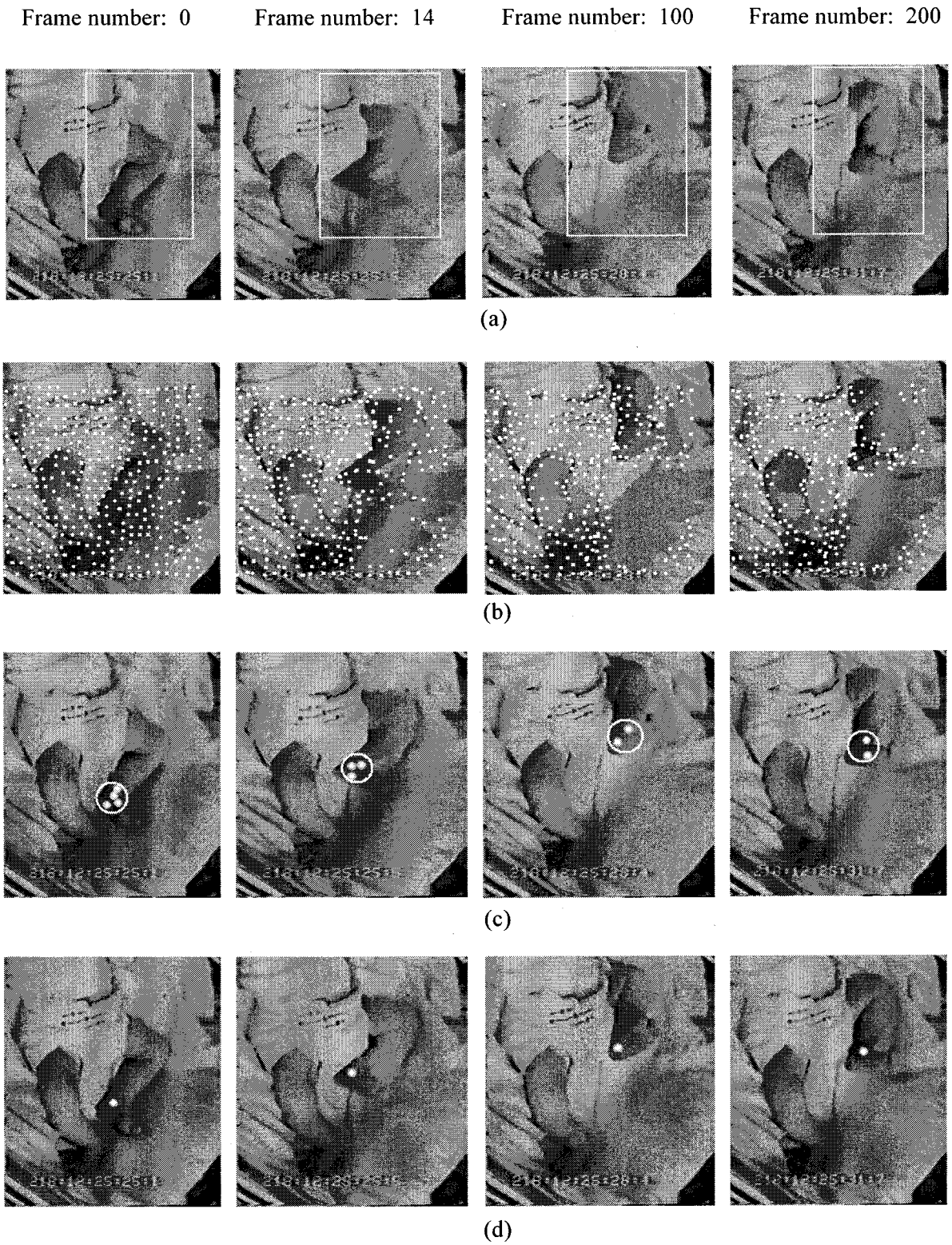

Fig. 4. Extraction of temporal motor activity signals. (a) Selected frames from a video-taped myoclonic seizure. (b) Frames containing the features tracked by the KLT algorithm. (c) Frames containing the features within a close neighborhood of the site. (d) Frames containing the feature whose projections to the horizontal and vertical axes produced the motor activity signals.

selected in each frame to produce the corresponding values of the temporal signals.

The experiments indicated that the selection of a small radius improved the accuracy of the tracking process, but also increased the likelihood of losing all features close to the site. The tradeoff typically associated with the selection of the radius motivated an extension of the KLT algorithm that improves its ability to recover the site even if all features in its close neighborhood are lost. In such a case, the location of the site in the next frame is predicted using the history of its motion in the

previous frames. Prediction was realized in this study through interpolation of the previous locations of the site. After the site was recovered by the method used for prediction, the KLT algorithm was used to track a new set of features in its close neighborhood.

V. Data Selection and Preprocessing

The two methods developed for extracting motion strength and motor activity signals from video recordings were tested and evaluated on video-taped clinical seizures selected from a database developed by the Clinical Research Centers for Neonatal Seizures (CRCNS). The CRCNS were established by the National Institute of Neurological Disorders and Stroke (NINDS) in 1991. The overall goal for this initiative was to develop a comprehensive understanding of the clinical and EEG features, predisposing risk factors, etiology, and outcome of seizures in the newborn. A comprehensive database has been created which includes detailed demographic information and maternal and infant risk factors, medical and neurological problems, neurological examinations, weekly tracking of subjects throughout hospitalization, and long-term follow-up at 6, 12, and 24 months of age. As part of this work, bedside video/EEG/polygraphic monitoring was performed (minimum of two hours for initial study), followed by repeat one-hour studies 3–5 days after the initial seizure characterization and at the time of discharge. Additional studies were performed whenever clinically indicated, particularly when new seizure behaviors occurred.

The CRCNS database contains several hundred individual clinical seizures, which are available to establish a library of motor signature patterns that are characteristic of focal clonic and myoclonic seizures in the newborn. Data from bedside video/EEG/polygraphic monitoring is available on videocassettes, including video recordings as well as digitized signals from EEG and polygraphic recordings. The seizures included in the CRCNS database have been characterized and classified by a team of clinical neurophysiologists and neonatal electroencephalographers in terms of their electrographic and behavioral features and the associated physiological manifestations have been documented. In making these determinations, the team members studied each video recording together with simultaneously recorded EEG. Decisions on characterization of seizures were made during group reviews (face-to-face discussions) in a way that a consensus was reached for each seizure included in the CRCNS database.

All seizures contained in the CRCNS database have been recorded in analog video. Analog-to-digital conversion of video involves the determination of the temporal and spatial sampling rates [43]. The temporal sampling rate specifies the number of frames that must be stored per second and depends mainly on the maximum velocity of the infant's moving body parts. The video-taped events were digitized in this study using a temporal sampling rate of 30 frames/s, which is typically used in applications requiring digital video of high temporal resolution. The spatial sampling rates specify the number of pixels that must be stored for each frame. In this particular application, the spatial sampling rates are not particularly important since the infants' upper and lower extremities occupy a large area of the video frames. Thus, the video-taped seizures selected from the available database were digitized to produce frames of the standard size of 352×240 pixels.

VI. Experimental Results

The two methods developed for extracting temporal signals from video recordings of neonatal seizures were tested on two myoclonic and two focal clonic seizures selected from the CRCNS database of neonatal seizures. Temporal signals were also produced for two video recordings of normal infant behavior (random infant movements). These video recordings are ordinarily used to test the ability of trained nurses to distinguish neonatal seizures from normal infant behaviors. The results of these experiments are summarized in Figs. 5–10, which show the four temporal signals extracted from each video-taped clinical event together with four representative frames of each sequence. These frames show the locations of the moving body parts during the clinical event and can be used to verify the consistency of the temporal signals with the clinical event captured by the video recording. The values of the signals corresponding to the frames shown at the bottom of each figure are indicated by dots, while the moving body part in each video sequence is shown within a box.

Figs. 5–10 indicate that there is an excellent correspondence between all four temporal signals extracted from each video recording and the motion of the body part of interest during the clinical event. For example, in the myoclonic seizure shown in Fig. 5, the infant's left leg moves to the right of the frame between frames 10 and 16 (Fig. 5 shows only frame 14). This movement is captured by the temporal signal obtained as the projection of the moving part to the horizontal axis. The temporal signal obtained as the projection of the moving part to the vertical axis indicates that the left leg also moves toward the top of the frame, which can be verified by comparing frames 0 and 14 of the sequence. The motor activity observed between frames 10 and 16 was also captured by the temporal signals measuring the area and scaled area of the moving part, as is clearly indicated by their spikes between frames 10 and 16. The infant's left leg remains at an almost fixed position between frames 50 and 140. In this time interval, the temporal motor activity signals are almost flat. The area and scaled area of the moving body part are almost zero between frames 50 and 140, which is consistent with the absence of any significant motion in this time interval. Right before frame 150, the left leg moves slightly to the left and toward the bottom of the frame, as is clearly indicated by the temporal motor activity signals. This movement is shown as a relatively weak spike in the temporal motion strength signals.

Inspection of Figs. 5–10 indicates that all four temporal signals obtained by the two methods developed in this study are consistent and reliable quantitative measures of the motor activity of the moving body part of interest. Moreover, the combination of all four temporal signals constitutes an effective representation of the clinical event captured by the video recording. Figs. 5–10 also indicate that the temporal signals produced by the two proposed approaches capture and quantify the differences between the motor activity of

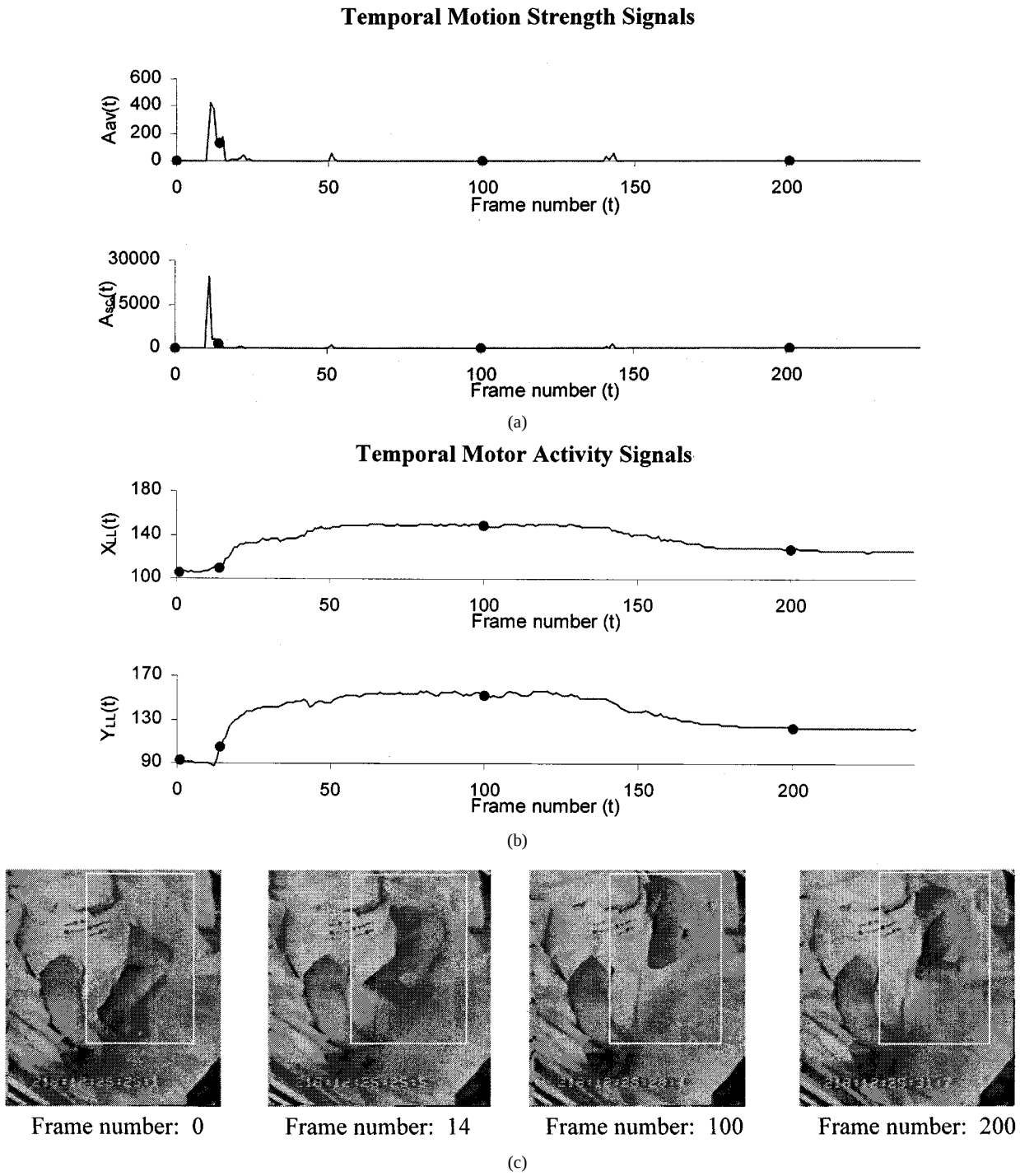

Fig. 5. Temporal signals produced for a video recording of a focal myoclonic seizure affecting the infant's left leg. (a) Temporal motion strength signals. (b) Temporal motor activity signals. (c) Selected frames of the sequence.

body parts caused by myoclonic and focal clonic seizures. In the case of myoclonic seizures, the temporal motor activity signals are consistent with the "jerky movements" that are the typical signatures of such events. The temporal motion strength signals contain a significant spike and a few weaker spikes. In the case of focal clonic seizures, the temporal motor activity signals capture and quantify the rhythmicity that characterizes the movements of such clinical events. The temporal motion strength signals contain multiple spikes that correspond very well with their rhythmic movements. The experiments on the focal clonic seizures also revealed the importance of the temporal signal representing the scaled area of the moving parts, that is, the signal obtained by multiplying the area of the moving parts by their displacement from frame to frame. This becomes obvious by inspecting Figs. 7 and 8, which indicate that some of the spikes obtained by measuring the area of the moving parts were reduced in magnitude when the scaled area was computed. This is consistent with the relatively low speed of the corresponding movements, which can also be observed from the temporal motor activity signals.

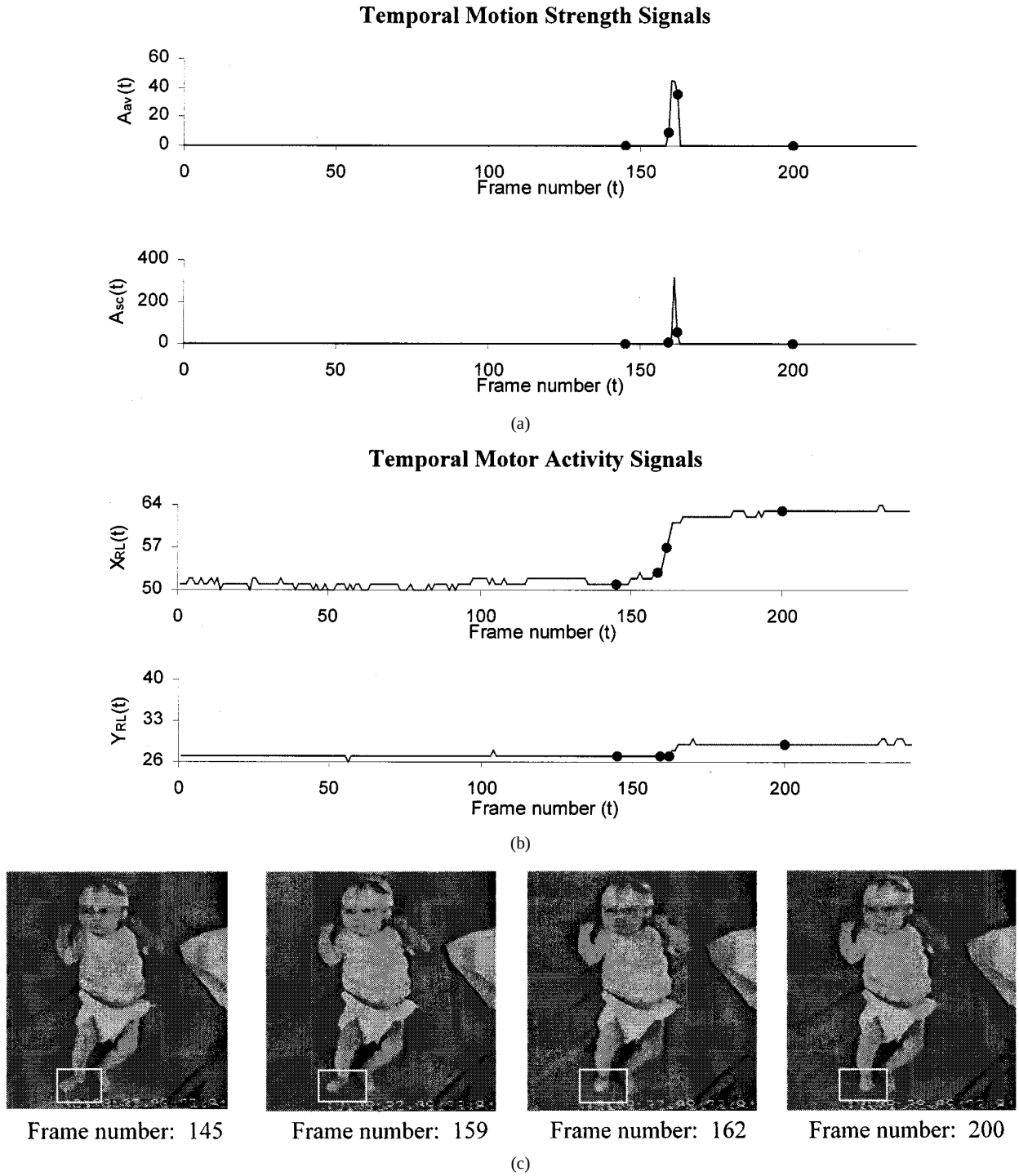

Fig. 6. Temporal signals produced for a video recording of a focal myoclonic seizure affecting the infant's right foot. (a) Temporal motion strength signals. (b) Temporal motor activity signals. (c) Selected frames of the sequence.

According to Figs. 5–10, the temporal signals produced by the two methods developed in this study provide a reliable basis for distinguishing normal infant behaviors from myoclonic and focal clonic seizures. The temporal motion strength signals produced for random movements of the infant's body parts contain fewer spikes compared with those corresponding to focal clonic seizures. The temporal motor activity signals produced for random movements of the infant's body parts contain bell-shaped spikes, which can easily be distinguished from the chain-saw-like signals produced for focal clonic seizures. There are also significant differences between the temporal signals produced for random movements of the infant's body parts and those corresponding to myoclonic seizures. The temporal motion strength signals produced for random movements of body parts contain spikes that are wider than those corresponding to myoclonic seizures. This experimental outcome is consistent with clinical observations, which indicated that random movements of body parts are typically slower than those caused by myoclonic seizures. This is also revealed by the rate at which temporal motor activity signals increase to

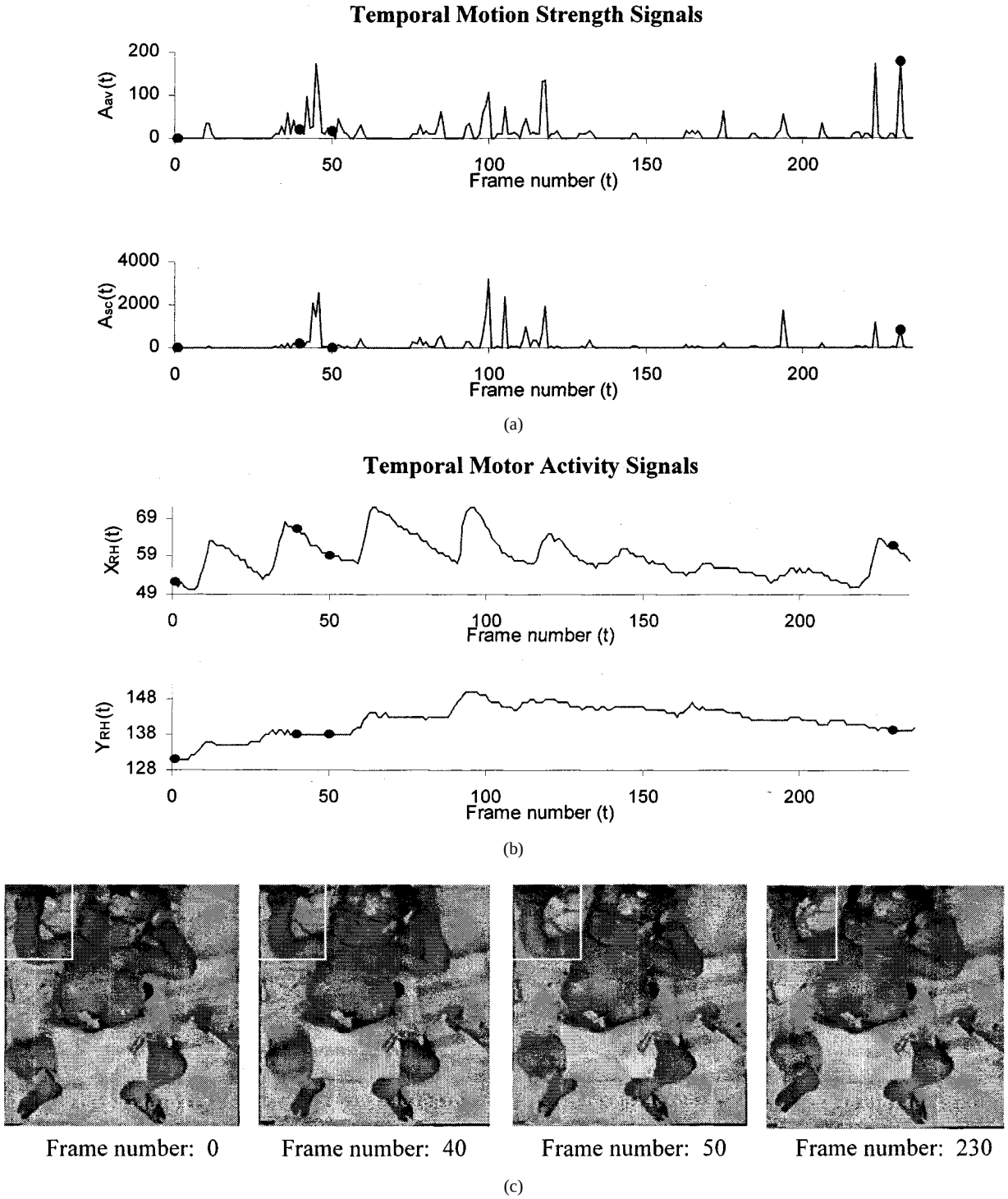

Fig. 7. Temporal signals produced for a video recording of a focal clonic seizure affecting the infant's right hand. (a) Temporal motion strength signals. (b) Temporal motor activity signals. (c) Selected frames of the sequence.

reach their peak value, which is lower in the case of random movements of the infant's body parts.

In conclusion, this experimental investigation indicated that the extraction of quantitative information from video-taped seizures in the form of temporal signals is feasible. According to the experimental results, the temporal signals extracted from video recordings of neonatal seizures provide a solid basis for selecting features that complement each other by conveying some unique behavioral characteristics of neonatal seizures. For example, the presence of motion of body parts as the clinical event progresses in time can be detected by computing the energy of motion strength signals in successive time intervals of the same duration. Myoclonic seizures and focal clonic seizures can be distinguished from nonseizure events, such as tremor and posturing of the extremities, by detecting the most significant spikes in the temporal motion strength signals. Short-time spectral analysis of motor activity signals can produce a set of features measuring the frequency of motion. Finally, the

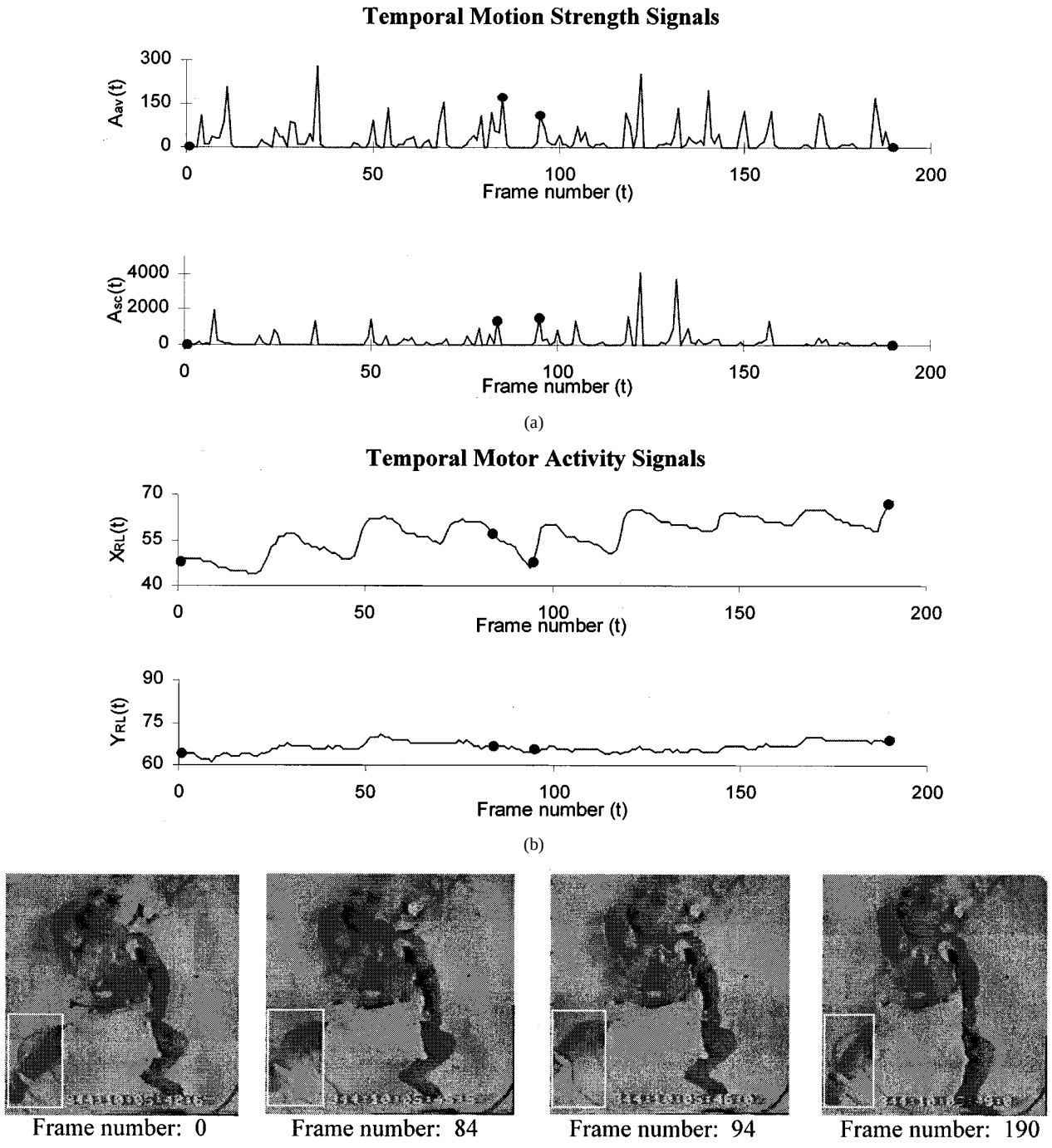

Fig. 8. Temporal signals produced for a video recording of a focal clonic seizure affecting the infant's right leg. (a) Temporal motion strength signals. (b) Temporal motor activity signals. (c) Selected frames of the sequence.

rhythmicity of movements can be quantified through features obtained from temporal motor activity signals by determining the locations and amplitudes of their peaks.

VII. Conclusion and Future Research

The temporal motion strength and motor activity signals extracted using the methods developed in this study can be used together with EEG monitoring to provide the basis for: 1) refining the characterization of repetitive motor behaviors; 2) improving the differentiation of certain clinical seizures from other abnormal paroxysmal behaviors not due to seizures; and 3) facilitating the detection of neonatal seizures. Even more importantly, the temporal signals extracted from video recordings can be utilized in the development of an automated system capable of recognizing focal clonic and myoclonic seizures and distinguishing them from clinical events characterized by high motor activity of the infants' extremities. Such a system could be developed by using a set of features computed in terms of the temporal motion strength and motor activity signals extracted from video to train classifiers based on artificial neural networks [2], [36].

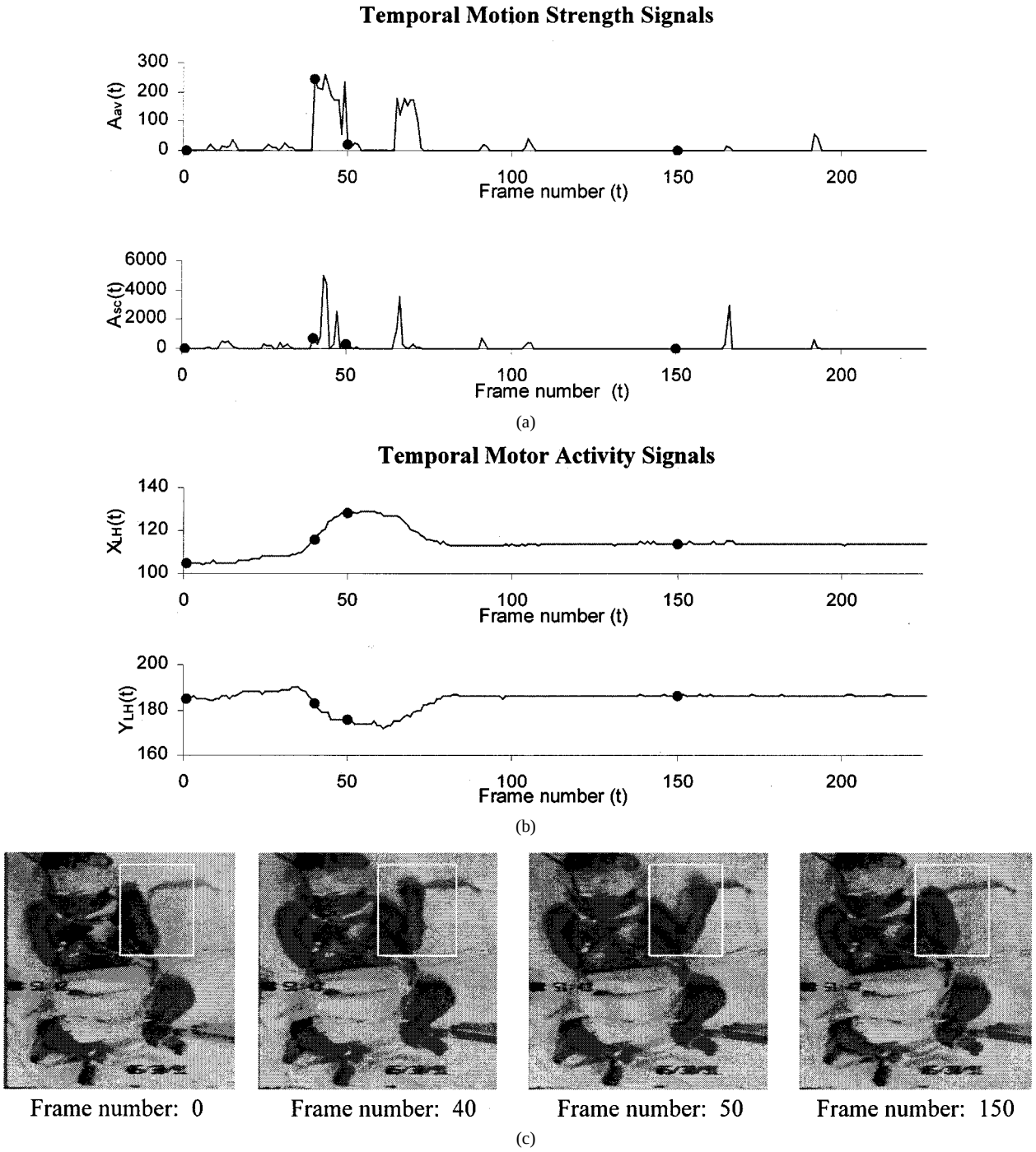

Fig. 9. Temporal signals produced for a video recording containing a random movement of the infant's left hand. (a) Temporal motion strength signals. (b) Temporal motor activity signals. (c) Selected frames of the sequence.

The methods described in this paper were tested on random infant movements and neonatal seizures of the focal clonic and myoclonic type. The proposed methods may not be suitable for all types of seizures, including those involving subtle movements of body parts other than the extremities. Ocular and orobuccolingual seizures are typical examples of these important types of clinical seizures [37], [45]. Nevertheless, focal clonic and myoclonic events constitute a large proportion of seizures observed in neonates in unselected populations [27], [29], [49]. Quantification of more "subtle" seizure types [45] or motor automatisms [29] may eventually be feasible with further development and refinement of these methods. The absence or "arrest" of ongoing motor activity can also be a manifestation of a seizure. Such cases can also be detected and quantified by extending and improving the methods described in this paper. Of course, these methods are not applicable to seizures whose only manifestation is electrical seizure activity in EEG [20]. The remainder of this section outlines some potential improvements in the design and implementation of the proposed methods.

The method used in this study to extract temporal motion strength signals relied on the segmentation of the filtered ver-

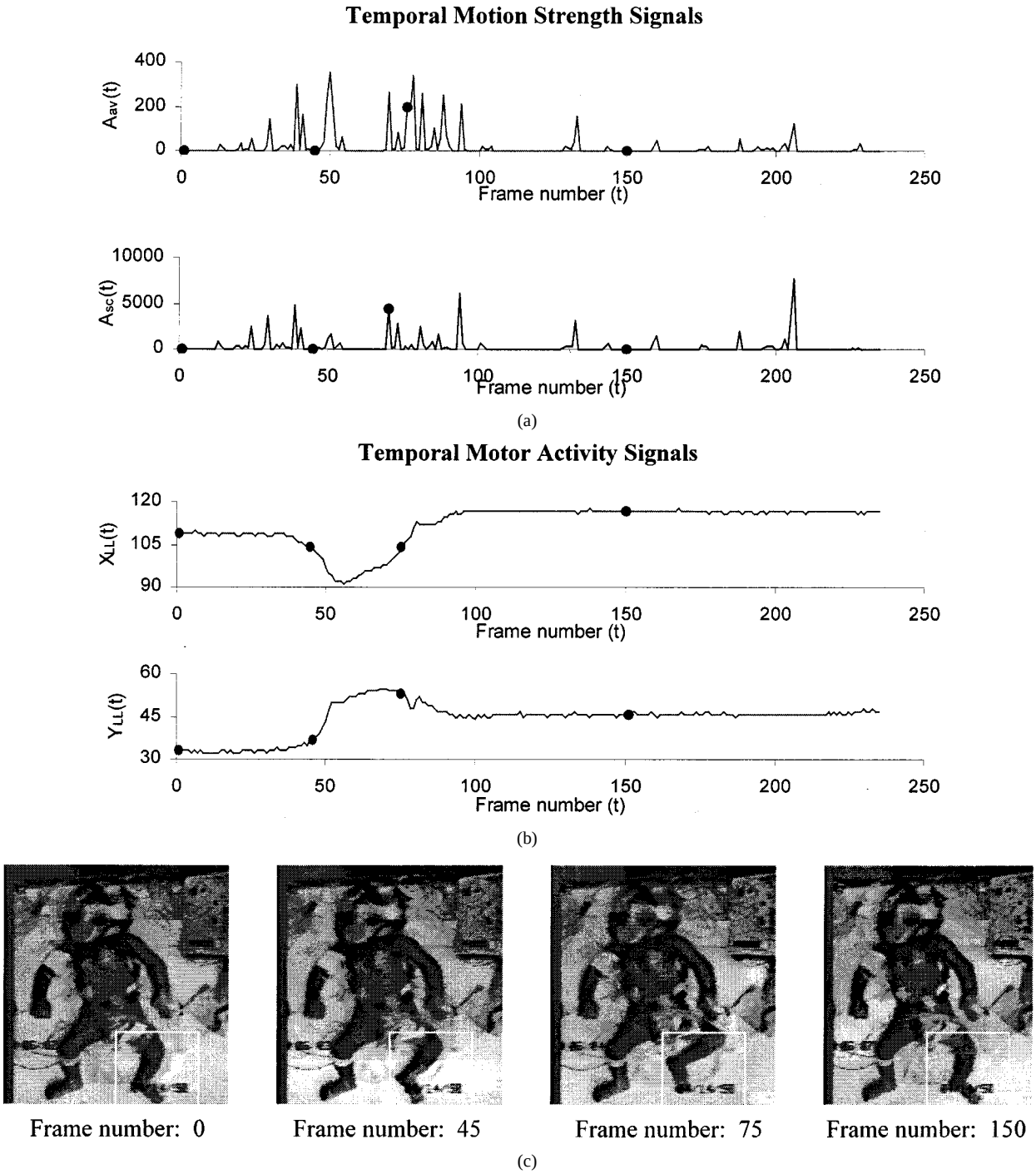

Fig. 10. Temporal signals produced for a video recording containing a random movement of the infant's left leg. (a) Temporal motion strength signals. (b) Temporal motor activity signals. (c) Selected frames of the sequence.

sion of subband 8. Segmentation was performed in this study by a clustering algorithm that was used to perform scalar quantization, which is naturally inferior to vector quantization. This indicates that the robustness of the signal extraction procedure may be improved by a segmentation technique that relies on vector quantization. In such a case, the clustering algorithm will be used to form clusters of vectors formed by groups of pixels. The results of segmentation may also be improved by replacing the c-means algorithm by competitive learning vector quantization algorithms, which have been successfully used to perform segmentation of magnetic resonance images of the brain [14]. Segmentation was performed in the experimental study by forming three clusters in order to accommodate the moving body parts, the background, and spurious clusters of pixels. This choice was motivated by the presence of spurious bright patches of intensities in between those of the background and the moving body parts. However, the formation of three clusters may create two clusters for the moving body parts if the level of noise in the frames is low. This can be prevented by selecting the number of clusters based on the signal-to-noise ratio computed for each frame sequence. The contribution of spurious patches to the measurements extracted from video recordings was prevented

in this study by tracking the centroids of the bright patches in the frame sequence produced by the segmentation process. Tracking was performed by considering only the areas whose centroids are present within a small radius between successive frames. This approach is effective if the motion between successive frames is within the predetermined radius of motion. This problem can be dealt with by developing an adaptive scheme that would modify the radius according to the motion present in the sequence. This can be accomplished by using as a reference the temporal motor activity signals extracted from the same frame sequence.

The method used in this study to extract temporal motor activity signals relied on the KLT algorithm. This algorithm was designed to select features that can be tracked through the entire frame sequence, but gives no priority to features located on moving objects in the sequence. This explains the loss of a substantial number of features located on the infant's moving parts during the seizure. In order to minimize the number of lost features affecting the extraction of motor activity signals, the KLT algorithm may be modified to track features located on moving objects with higher priority. This can be accomplished by including motion in the criteria used for rejecting features during the sequence. This improvement is expected to reduce the rejection of features located on moving body parts, but it is unlikely that this problem will be completely eliminated. This implies that the automated procedure used to extract temporal motor activity signals may still stop tracking a selected anatomical site in a frame sequence if all features in its close neighborhood are lost. This reveals the need for a reliable procedure that can be used to predict the location of the site in the next frame if the algorithm stops tracking the features within its close neighborhood in the current frame. Prediction was performed in this study by simple interpolation, which may not be always reliable. The reliability of prediction can be improved by applying optimal linear prediction techniques or nonlinear prediction models based on neural networks. The automated procedure developed in this study for extracting temporal motor activity signals is capable of tracking a single site throughout each frame sequence. However, the CRCNS database of video-taped neonatal seizures contains events involving movements of multiple body parts. The extraction of temporal motor activity signals for such events can be accomplished by extending and improving the existing procedure to make it capable of tracking multiple sites throughout each frame sequence. Tracking of multiple anatomical sites during the seizure also requires the development of an automated procedure for detecting the moving body part(s).

REFERENCES

[1] C. D. Binnie, A. J. Rowan, J. Overweg, H. Meinard, T. Wisman, A. Kamp, and F. L. da Silva, "Telemetric EEG and video monitoring in epilepsy," *Neurology*, vol. 31, pp. 298–303, 1981.

[2] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press,, 1995.

[3] A. Bye, P. Lamont, and L. Healy, "Commencement of a pediatric EEG-video telemetry service," *Clin. Exp. Neurol.*, vol. 27, pp. 83–88, 1990.

[4] A. Bye and D. Flanagan, "Electroencephalograms, clinical observations and the monitoring of neonatal seizures," *J. Paediatrics Child Health*, vol. 31, no. 6, pp. 503–507, 1995.

[5] A. M. Bye and D. Flanagan, "Spatial and temporal characteristics of neonatal seizures," *Epilepsia*, vol. 36, no. 10, pp. 1009–1016, 1995.

[6] A. Cappozzo, T. Leo, and A. Pedotti, "A general computing method for the analysis of human locomotion," *J. Biomech.*, vol. 8, pp. 307–320, 1975.

[7] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[8] J. D. Frost Jr., R. A. Hrachovy, P. Kellaway, and T. Zion, "Quantitative analysis and characterization of infantile spasms," *Epilepsia*, vol. 19, pp. 273–282, 1978.

[9] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston: Kluwer Academic, 1992.

[10] J. R. Ives and P. Gloor, "A long term time-lapse video system to document the patient's spontaneous clinical seizure synchronized with the EEG," *Electroencephalogr. Clin. Neurophysiol.*, vol. 45, pp. 412–416, 1978.

[11] J. R. Ives, N. R. Mainwaring, L. J. Gruber, G. R. Cosgrove, H. W. Blume, and D. L. Schomer, "128-channel cable-telemetry EEG recording system for long-term invasive monitoring," *Electroencephalogr. Clin. Neurophysiol.*, vol. 79, pp. 69–72, 1991.

[12] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[13] N. B. Karayiannis and Y. Li, "A replenishment technique for low bit-rate video compression based on wavelets and vector quantization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 658–663, May 2001.

[14] N. B. Karayiannis and P.-I. Pai, "Segmentation of magnetic resonance images using fuzzy algorithms for learning vector quantization," *IEEE Trans. Med. Imag.*, vol. 18, pp. 172–180, Feb. 1999.

[15] N. B. Karayiannis, P.-I. Pai, and N. Zervos, "Image compression based on fuzzy algorithms for learning vector quantization and wavelet image decomposition," *IEEE Trans. Image Processing*, vol. 7, pp. 1223–1230, Aug. 1998.

[16] N. B. Karayiannis and T. C. Wang, "Compression of digital mammograms using wavelets and fuzzy algorithms for learning vector quantization," in *Soft Computing for Image Processing*, S. K. Pal, A. Ghosh, and M. K. Kundu, Eds, Heidelberg, Germany: Physical-Verlag, 2000, pp. 205–245.

[17] P. Kellaway and J. D. Frost Jr., "Monitoring at the Baylor College of Medicine, Houston," in *Long-term Monitoring in Epilepsy*, J. Gotman, J. R. Ives, and P. Gloor, Eds. Amsterdam, The Netherlands: Elsevier Science, 1985, pp. 403–414.

[18] P. Kellaway, R. A. Hrachovy, J. D. Frost Jr., and T. Zion, "Precise characterization and quantification of infantile spasms," *Ann. Neurol.*, vol. 6, no. 3, pp. 214–218, 1978.

[19] H. D. Kim and R. R. Clancy, "Sensitivity of a seizure activity detection computer in childhood video/electroencephalographic monitoring," *Epilepsia*, vol. 38, no. 11, pp. 1192–1197, 1997.

[20] N. Laroia, R. Guillet, J. Burchfiel, and M. C. McBride, "EEG background as predictor of electrographic seizures in high-risk neonates," *Epilepsia*, vol. 39, no. 5, pp. 545–551, 1998.

[21] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereoscopic vision," in *Proc. Int. Conf. Artificial Intelligence*, 1981, pp. 674–679.

[22] J. S. Luther, J. O. McNamara, S. Carwile, P. Miller, and V. Hope, "Pseudoepileptic seizures: Methods and video analysis to aid diagnosis," *Ann. Neurol.*, vol. 12, no. 5, pp. 458–462, 1982.

[23] S. G. Mallat, "Multifrequency channel decomposition of images and wavelet models," *IEEE Trans. Acoust., Speech Signal Processing*, vol. 37, pp. 2091–2110, Dec. 1989.

[24] M. C. McBride, N. Laroia, and R. Guillet, "Electrographic seizures in neonates correlate with poor neurodevelopmental outcome," *Neurology*, vol. 55, no. 4, pp. 506–513, 2000.

[25] E. M. Mizrahi, "Neonatal electroencephalography: Clinical features of the newborn, techniques of recording and characteristics of the normal EEG," *Amer. J. EEG Technol.*, vol. 26, pp. 81–103, 1986.

[26] ——, "Neonatal seizures," in *Pediatric and Adolescent Medicine*, S. Shinnar, N. Amir, and D. Branski, Eds. Basel, Switzerland: Karger, vol. 6, pp. 18–31.

[27] E. M. Mizrahi, R. R. Clancy, J. K. Dunn, D. G. Hirtz, L. Chapieski, S. A. McGaurn, P. Cuccaro, R. A. Hrachovy, M. S. Wise, and P. Kellaway, "Neurological impairment, developmental delay and postneonatal seizures two after EEG-video documented seizures in near-term and full-term neonates: Report of the clinical research center for neonatal seizures," *Epilepsia*, to be published.

[28] E. M. Mizrahi and P. Kellaway, "Characterization of seizures in neonates and young infants by time-synchronized electroencephalographic/polygraphic/video monitoring," *Ann. Neurol.*, vol. 16, p. 383, 1984.

[29] ——, "Characterization and classification of neonatal seizures," *Neurology*, vol. 37, pp. 1837–1844, 1987.
[30] M. E. Morris, T. A. Matyas, R. Iansek, and J. J. Summers, "Temporal stability of gait in Parkinson's disease," *Physical Ther.*, vol. 76, no. 7, pp. 763–789, 1996.
[31] H. Oguni, Y. Fukuyama, Y. Imaizumi, and T. Uehara, "Video-EEG analysis of drop seizures in myoclonic astatic epilepsy of early childhood (Doose syndrome)," *Epilepsia*, vol. 33, no. 5, pp. 805–813, 1992.
[32] J. K. Penry, R. J. Porter, and F. E. Dreifuss, "Simultaneous recording of absence seizures with video tape and electroencephalography," *Brain*, vol. 98, pp. 427–440, 1975.
[33] F. Pierelli, G.-E. Chatrian, W. W. Erdly, and P. D. Swanson, "Long-term EEG-video-audio monitoring: Detection of partial epileptic seizures and psychogenic episodes by 24-hour EEG record review," *Epilepsia*, vol. 30, no. 5, pp. 513–523, 1989.
[34] C. Podilchuk and A. Jacquin, "Subband video coding with a dynamic bit allocation and geometric vector quantization," in *SPIE Proc.*, vol. 1666, 1992, pp. 241–252.
[35] D. Rector, P. Burk, and R. M. Harper, "A data acquisition system for long-term monitoring of physiological and video signals," *Electroencephalogr. Clin. Neurophysiol.*, vol. 87, pp. 380–384, 1993.
[36] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
[37] M. S. Scher, K. Aso, M. Beggarly, M. Y. Hamid, D. A. Steppe, and M. J. Painter, "Electrographic seizures in preterm and full-term neonates: Clinical correlates, associated brain lesions and risk for neurologic sequelae," *Pediatrics*, vol. 91, no. 1, pp. 128–134, 1993.
[38] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1994, pp. 593–600.
[39] S. Srinivasan, R. Bhattacharya, N. B. Karayiannis, M. S. Wise, J. D. Frost Jr., and E. M. Mizrahi, "Extraction of motion strength and motor activity signals from video recordings of neonatal seizures," in *Proc. 18th Annu. Houston Conf. Biomedical Engineering Research*, Houston, TX, Feb. 10–11, 2000, p. 171.
[40] V. P. Stokes, "A method for obtaining the 3-D kinematics of the pelvis and thorax during locomotion," *Human Movement Sci.*, vol. 3, pp. 77–94, 1984.
[41] V. P. Stokes, C. Anderson, and H. Forssberg, "Rotational and translational movement features of the pelvis and thorax during adult human locomotion," *J. Biomech.*, vol. 22, pp. 43–50, 1989.
[42] V. P. Stokes, H. Lanshammar, and A. Thorstensson, "Dominant pattern extraction from 3-D kinematic data," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 100–106, Jan. 1999.
[43] A. M. Tekalp, *Digital Video Processing*. Englewood Cliffs, NJ: Prentice Hall, 1995.
[44] C. Tomasi and T. Kanade, "Detection and Tracking of Point Features," Carnegie Mellon Univ. , Pittsburgh, Tech. Rep. CMU-CS–91–132 , April 1991.
[45] J. J. Volpe, *Neurology of the Newborn*, 4th ed. Philadelphia, PA: Saunders, 2000.
[46] T. C. Wang and N. B. Karayiannis, "Detection of microcalcifications in digital mammograms using wavelets," *IEEE Trans. Med. Imag.*, vol. 17, pp. 498–509, Aug. 1998.
[47] S. P. Weiner, M. J. Painter, D. Geva, R. D. Guthrie, and M. S. Scher, "Neonatal seizures: Electroclinical dissociation," *Pediatric Neurol.*, vol. 7, no. 5, pp. 363–368, 1991.
[48] D. A. Winter, A. O. Quanbury, D. A. Hobson, H. G. Sidwall, G. Reiner, B. G. Trenholm, T. Steinke, and H. Shlosser, "Kinematics of normal locomotion—A statistical study based on T.V. data," *J. Biomech.*, vol. 7, pp. 479–486, 1974.
[49] M. S. Wise, E. M. Mizrahi, R. A. Hrachovy, R. R. Clancy, J. K. Dunn, J. Lane, S. A. McGaurn, and D. G. Hirtz, "Seizures in very low birthweight (VLBW) infants: Seizure characterization using bedside EEG/video/polygraphic monitoring," *Epilepsia*, vol. 40 (Suppl. 7), pp. 161–161, 1999.
[50] M. Y. Zarrugh and C. W. Radcliff, "Computer generation of human gait kinematics," *J. Biomech.*, vol. 12, pp. 99–111, 1979.