Improving the Extraction of Temporal Motion Strength Signals From Video Recordings of Neonatal Seizures

Nicolaos B. Karayiannis and Guozhi Tao
Department of Electrical and Computer Engineering
N308 Engineering Building 1
University of Houston
Houston, TX 77204-4005
Karayiannis@uh.edu, Guozhi.Tao@mail.uh.edu

Abstract

This paper presents a procedure developed to extract quantitative information from video recordings of neonatal seizures in the form of temporal motion strength signals. These signals are obtained by applying nonlinear filtering, segmentation, and morphological filtering on the differences between adjacent frames. The experiments indicate that temporal motion strength signals constitute an effective representation of videotaped clinical events and can be used for seizure recognition and characterization.

1. Introduction

Seizures occur in approximately 2-5/1000 live births, depending upon studied populations and methodology [4], [9], [12]. In fact, the incidence of seizures in infants weighing less than 1500 grams is 57.5/1000 live births compared to 3.5/1000 live births for all birthweights [9]. Similarly, seizures occurred in approximately 4% of premature infants less than 30 weeks conceptional age, although some have reported the incidence in this population reaching as high as 20% [13]. These studies indicate that seizure occurrence represents the most frequent clinical sign of central nervous system disorders in the newborn [2], [11], [14]. These disturbances in cerebral function may result in significant long-term adverse sequelae such as neurological handicaps, mental retardation, and postnatal epilepsy [1], [4], [7], [10], [14]. The prompt identification of clinical seizures when they occur in the newborn, the subsequent evaluation of their etiology, and the institution of etiology-specific therapy may significantly reduce associated morbidity.

Automated processing and analysis of video recordings of neonatal seizures can generate novel methods for extracting quantitative information that is relevant only to the seizure [5], [6]. This information can be used to: 1) develop automated mechanisms capable of detecting the beginning of clinical seizures, 2) refine the characterization of repetitive motor behaviors, and 3) facilitate the differentiation of certain clinical seizures from other abnormal paroxysmal behaviors not due to seizures. The development of an automated video analysis system would represent a major advance in seizure surveillance and offers the possibility for earlier identification of potential neurological problems and subsequent intervention.

This paper summarizes the results of a study that aimed at improving the extraction of motion strength signals from video recordings of neonatal seizures. In principle, motion strength signals quantify motion by measuring the area of the frames occupied by moving body parts affected by seizures.

2. Extraction of temporal motion strength signals from video

The extraction of quantitative information from videotaped seizures must focus only on the moving parts of the infant's body that are affected by the seizure [5], [6]. The extraction from video recordings of visual information that is relevant only to the seizure can be accomplished by quantifying motion between adjacent frames of the video recording of a neonatal seizure.

2.1 Detection of motion

In this study, motion was detected on a new frame sequence obtained by subtracting the current frame from the previous frame. The experiments indicated that the difference frames contain the moving body parts but they are 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. Removing random spiky noise is essential since this kind of noise can mislead the clustering algorithm used in the next stage for segmentation. Spiky noise was removed in this study by applying median filtering, which is very effective in removing this type of noise without blurring the edges of the objects [3].

Figure 1(b) shows the frames resulting from the application of a 2-D median filter of size 3×3 pixels on the difference frames corresponding to the videotaped focal clonic seizure shown in Figure 1(a). The difference frames corresponding to frames 82, 150, and 190 show the infant's right leg, which is affected by the seizure. In frame 82, the right leg moves to the bottom and to the left of the frame. In frame 150, the right leg moves to the top and to the right of the frame.

2.2 Segmentation

Following median filtering, the time sequence formed by the difference frames was segmented in order to isolate the moving body parts from background noise and other irrelevant clusters of pixels. Segmentation was performed in the original study by an adaptive version of the c-means (or k-means) algorithm, which clustered all pixels of each frame from the sequence formed by the difference frames into $c = 3$ clusters [5], [6]. The main drawback of such a segmentation procedure is that it ignores the spatial correlation within each frame since it forms clusters of individual pixels. Moreover, the use of scalar clustering degrades the resistance of the segmentation procedure to noise. The robustness of the procedure developed to extract motion strength signals was improved by employing a segmentation technique that relies on clustering of vectors instead of scalar clustering. More specifically, the c-means algorithm was used to form $c = 4$ clusters of vectors formed by groups of neighboring pixels. This procedure clustered vectors of length 9 that contained the pixels of a 3×3 window.

The pixels of the difference frames with nonzero values belong to a region of the frame occupied by the moving body part. Note that the infant's body and the background have different intensity values in the original frames. Thus, if an object moves from a certain position in the previous frame to a new position in the current frame, the intensity differences in the old and new positions of the object will have different signs. In the original study, the pixels corresponding to the old and new positions were segmented together [5], [6]. In the approach described in this paper, the pixels corresponding to the old and new positions of the moving object formed two different segments. This was accomplished by employing the c-means algorithm to partition the available vectors into $c = 4$ clusters. The c-means algorithm was initialized to produce a cluster corresponding to pixels of positive intensity values of large magnitude, a cluster corresponding to pixels of negative intensity values of large magnitude, and a cluster containing background pixels. The fourth cluster was introduced deliberately in order to accommodate pixels that do not belong to the background but they are not relevant to the seizure. Following cluster formation, 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). The pixels belonging to the cluster of the lowest intensities were all assigned the same intensity value of 0 (corresponding to black color). Finally, the pixels corresponding to the other two clusters were all considered background and were assigned the same intensity value of 128 (corresponding to gray color).

Figure 1(c) shows the difference frames corresponding to the four frames in Figure 1(a) following segmentation. Segmentation eliminated most spurious clusters of pixels in frame 190. Segmentation also eliminated the clusters of pixels with low absolute intensity values from frame 82 and 150, which led to a better definition of the moving body part. The traces of the infant's left leg are shown in frame 82 and 150 as white and black patches in a gray background. It is clear from Figure 1(c) that the two black and white segments correspond well with the motion of the right leg, which moves along the horizontal direction. Moreover, the relative locations of the black and white segments reveal the direction of motion. Nevertheless, the segmented frames shown in Figure 1(c) contain some spurious black and white patches of relatively small sizes. These patches are rather isolated and do not correspond to the moving body part. These patches can be eliminated, or at least reduced, by applying nonlinear binary operators developed for mathematical morphology [3], [8].

2.3 Morphological filtering

The frames produced by the segmentation process outlined above still contained spurious white and black patches (i.e., small groups of pixels with high and low intensity values relative to the gray background). Such patches do not typically belong to body parts affected by the seizure and they resemble 'salt-and-pepper' noise. The reduction of such spurious patches was attempted by relying on the OPENING and CLOSING morphological operations [3], [8]. The OPENING of an object A by a structuring element X, denoted as $A \circ X$, is the erosion of A by X followed by a dilation of the result by X. In mathematical terms $A \circ X = (A \ominus X) \oplus X$, where $A \ominus X$ and $A \oplus X$ denote the erosion and dilation of the object

A by the structuring element X, respectively. OPENING smoothes the contour of an object and breaks narrow bridges. The CLOSING of an object A by a structuring element X, denoted as $A \bullet X$, is the dilation of A by X followed by an erosion of the result by X. In mathematical terms, $A \bullet X = (A \oplus X) \odot X$. CLOSING also smoothes the contour of an object. However, in contrast with OPENING, CLOSING fuses narrow breaks, eliminates small holes, and fills gaps in the contour.

In this study, each of the segmented frames (containing three intensity levels) produced two binary frames. One of the binary frames contained only the white patches of the segmented frame while the rest of the frame (including the black patches) was considered to be background. The other binary frame contained only the black patches of the segmented frame while the rest of the frame (including the white patches) was considered to be background. The spurious black and white patches were reduced by applying a sequence of morphological operators to the binary frames. More specifically, each of the two binary frames was operated first by the OPENING morphological operator followed by the CLOSING morphological operator. Both operators employed the same structuring element, which was selected to be a circle within a 5×5 square window. The OPENING operator reduced the size of the big black and white patches corresponding to the moving body parts and eliminated the black and white patches that were relatively small in size and isolated from each other. The CLOSING operator that followed restored the size and shape of the black and white patches that belong to body parts affected by the seizure. The binary frames resulting after the application of these two morphological operators were subsequently used to restore the segmented frame. This was accomplished by copying the black and white patches left intact after the morphological operators and considering the rest of the frame to be background.

Figure 1(d) shows that the frames produced by applying morphological filtering on the segmented frames shown in Figure 1(c). It is clear that morphological filtering eliminated most of the noisy black and white patches that were not located at the infant's moving body part. Motion of the infant's right leg was clearly identified in frames 82 and 150. Comparison of Figures 1(c) and 1(d) indicates that the size of the black and white patches generated by the motion of the infant's right leg was reduced slightly after the application of the OPENING and CLOSING morphological operators. The location of the black and white patches in Figure 1(d) clearly reveals the direction of motion. Finally, morphological filtering eliminated most of the noisy black and white patches that can be seen in frame 190 of Figure 1(c).

2.4 Extraction of motion strength signals

The motion strength signals were obtained by measuring the area of the moving body parts. In order to improve the robustness of the proposed procedure to noise, the extraction of motion strength signal exploited the black and white patches produced by the segmentation process and refined by morphological filtering. Note that the black and white patches represent the areas occupied by the moving body part in two successive frames of the sequence. Since the black and white patches represent motion of the same body part between successive frames, their areas are expected to be equal or, at least, sufficiently close to each other. In the presence of noise, the area occupied by either the black patches or the white patches may be abnormally high. In order to reduce the sensitivity of the proposed procedure to noise, the area of the moving body part was calculated as the minimum of the areas occupied by black patches and white patches.

3. Experimental results

The two procedures described above were evaluated on the video recordings of neonatal seizures and other clinical events representing normal and abnormal infant behaviors not due to seizures [6]. Figures 2, 3 and 4 show the temporal motion strength signals extracted from video recordings of a myoclonic seizure, a focal clonic seizure, and a random infant movement, respectively. The value of the motion strength signal $A(t)$ at each frame t of the video recording represents the number of pixels from the area of the frame identified by the proposed procedure as belonging to the moving body part. The locations of the moving body parts during the clinical event are shown in representative frames of each video recording. The frames of the video recordings shown in Figures 2-4 can be used as a reference to verify the consistency of the temporal signals with the corresponding clinical events. The values of the signals corresponding to the frames shown at the top of each figure are indicated by dots, while the moving body part in each video recording is shown within a box.

In the myoclonic seizure shown in Figure 2, the infant's left leg moves to the right of the frame between frames 10 and 16 (Figure 2 shows only frame 14). Both approaches identified that the left leg moved between frames 10 and 16, which is indicated by the sizable spike in this time interval shown in Figures 2(c) and 2(d). However, the proposed procedure produced a higher spike; this indicates that, according to the proposed method, the moving body part occupied a larger area of the frames.

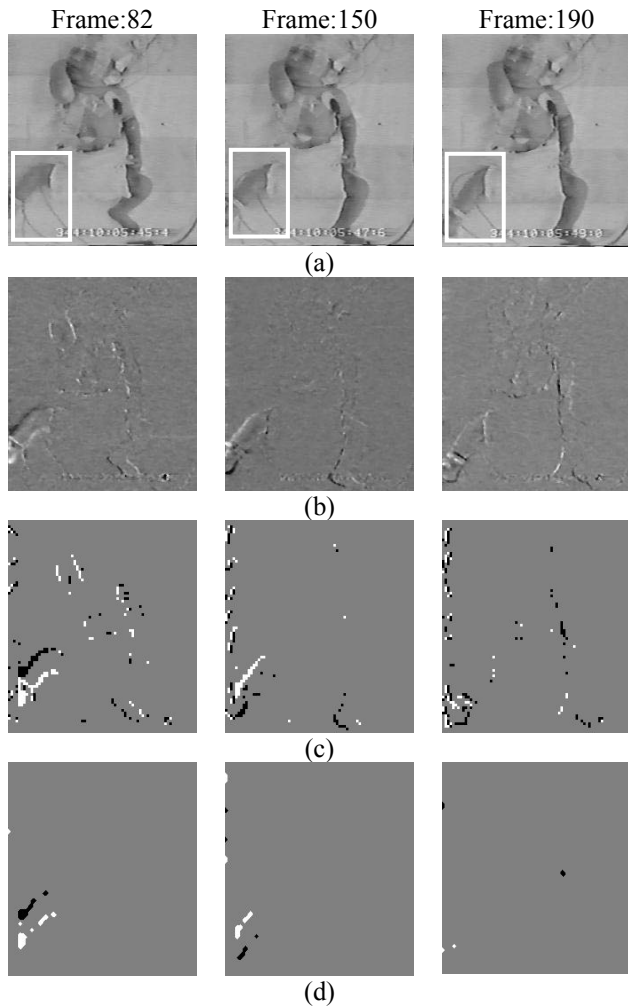| Frame:82 | Frame:150 | Frame:190 |
| Frame:3 | Frame:14 | Frame:16 |

(a)

(b)

(c)

(d)

Figure 1: Extraction of temporal motion strength signals: (a) selected frames from a video taped focal clonic seizure, (b) frames produced by applying median filtering on the difference frames, (c) frames produced by segmenting the filtered difference frames and (d) frames produced by morphological filtering.
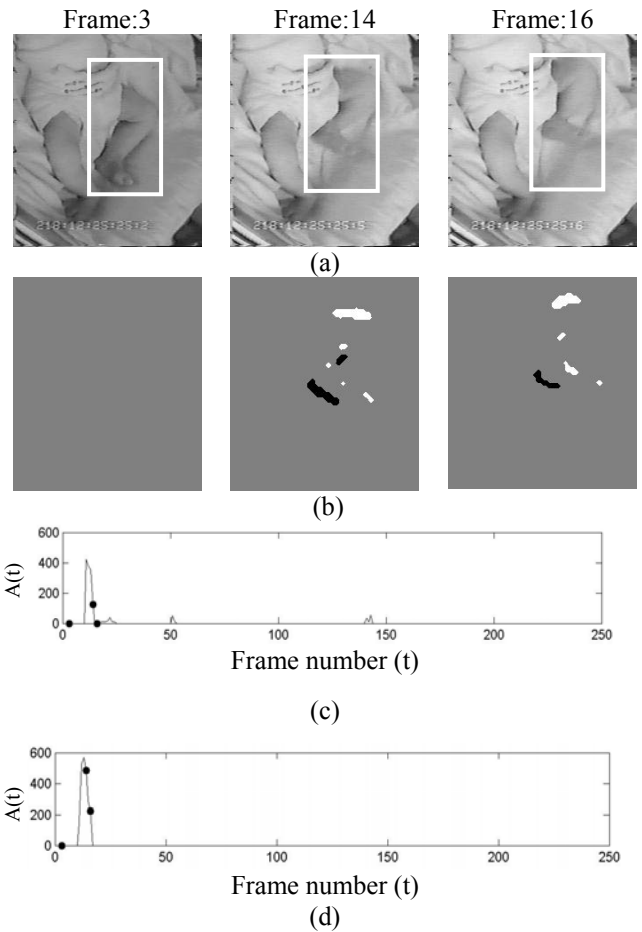

(a)

(b)

(c)

(d)

Figure 2: Temporal signals produced for a video recording of a myoclonic seizure affecting the infant's left leg: (a) selected frames of the sequence, (b) difference frames following segmentation and morphological filtering, (c) temporal motion strength signal produced by the original procedure, and (d) temporal motion strength signal produced by the proposed procedure.

According to Figure 2(d), the proposed method identified no motion after frame 16. In contrast, the motion strength signal produced by the original procedure contains some weaker spikes after the big spike between frames 10 and 16. It was visually verified that the infants left leg remains at an almost fixed position after frame 16. Thus, the weak spikes shown in Figure 2(c) can be attributed to noise that interferes with the extraction of the motion strength signal or to some weak motion that is not visually detectable when reviewing the video recording frame-by-frame. Nevertheless, even if the weak spikes in Figure 2(c) represent motion, this motion is too weak to be attributed to the myoclonic seizure. Note that the signature

characteristic of a myoclonic seizure is the rapid motion of the extremities affected by the seizure. The temporal motion strength signals shown in Figure 2 contain a sharp spike that corresponds to strong motion due to the seizure. Thus, both signals shown in Figure 2 are consistent with the "jerky" movements that are the typical signatures of myoclonic seizures.

The motion strength signals produced for the focal clonic seizure shown in Figure 3(a) reveal the rhythmicity that is the signature characteristic of this clinical event. Comparison of Figures 3(c) and 3(d)

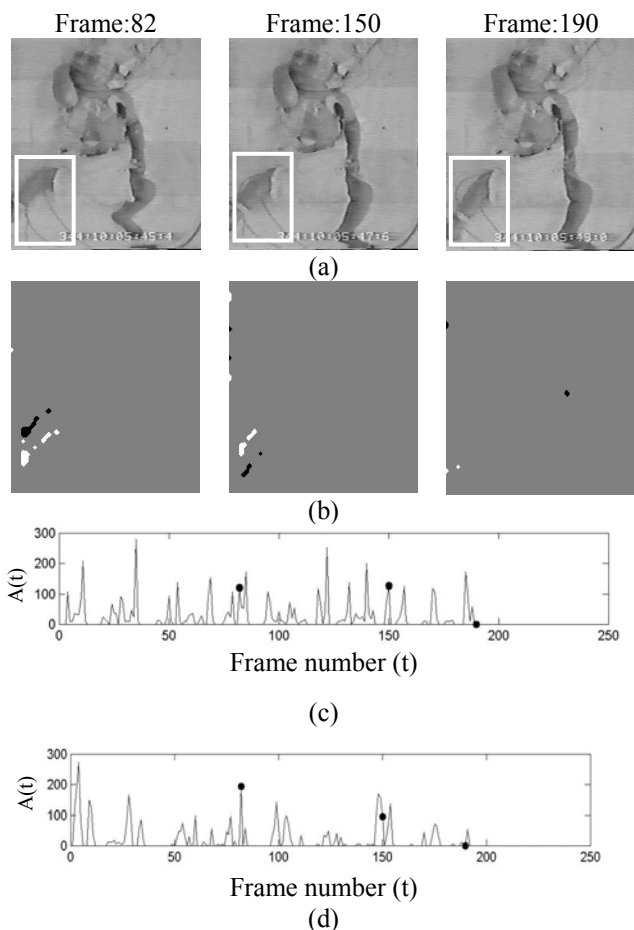| Frame:82	Frame:150	Frame:190

(a)

(b)

(c)

(d)

Figure 3: Temporal signals produced for a video recording of a focal clonic seizure affecting the infant's right leg: (a) selected frames of the sequence, (b) difference frames following segmentation and morphological filtering, (c) temporal motion strength signal produced by the original procedure, and (d) temporal motion strength signal produced by the proposed procedure.

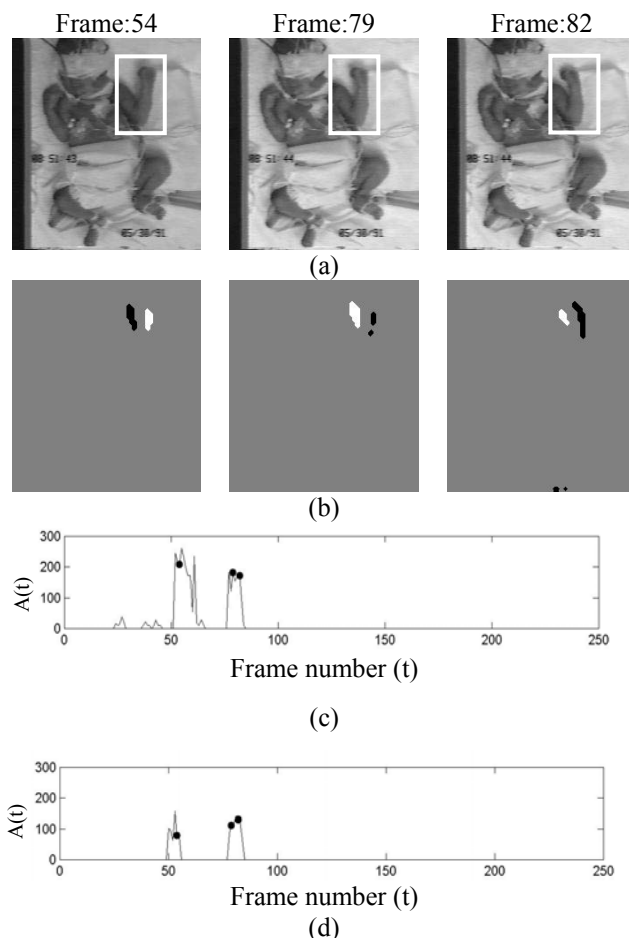| Frame:54	Frame:79	Frame:82

(a)

(b)

(c)

(d)

Figure 4: Temporal signals produced for a video recording of a random movement of the infant's left arm: (a) selected frames of the sequence, (b) difference frames following segmentation and morphological filtering, (c) temporal motion strength signal produced by the original procedure, and (d) temporal motion strength signal produced by the proposed procedure.

indicates that the two procedures employed in the experiments identified motion at the same frames of the video recording. This becomes clear by the locations of the spikes produced by the two procedures. However, there are some noteworthy differences between the motion strength signals produced by the two procedures. The original procedure identified strong motion between frames 100 and 150. In contrast, the procedure proposed in this paper identified only weak motion in this time interval. On the other hand, the proposed procedure identified stronger motion right before and after frame 100. Compared with the original method, similar differences can also be seen between frames 0 and 50. Overall, the experiments indicated that the temporal signals produced by the

proposed procedure capture and quantify the differences between myoclonic and focal clonic seizures. In the case of focal clonic seizures, the temporal motion strength signals contain multiple spikes that correspond well with the rhythmicity that characterizes the movements of such clinical events.

Figure 4 shows the temporal motion strength signals produced by the original and the proposed procedures for the video recording of a random movement of an infant's left arm. The two procedures produced spikes located at almost the same frames of the video recording as indicated by comparing Figures 4(c) and 4(d).

These spikes represent the motion strength of the moving body part. As such, they are not as sharp as those produced for myoclonic seizures. This is consistent with the motion patterns that differentiate random movements from the rapid movements caused by myoclonic seizures.

4. Conclusions

This paper outlined a new procedure proposed for extracting temporal motion strength signals from video recordings of neonatal seizures. This procedure employs nonlinear filtering, segmentation based on vector clustering, and morphological filtering. The outcome of the experiments indicates that the proposed method is less susceptible to noise than the original method. As a result, the proposed method typically produces fewer significant spikes than the original procedure. On the other hand, it is possible that the proposed procedure underestimates the area occupied by the moving body part in an attempt to eliminate the "salt-and-pepper" noise present in the segmented difference frames.

Further improvement and refinement of the procedure developed in this study can produce temporal motion strength signals that constitute a consistent and effective representation of videotaped clinical events. The trade-off between noise reduction and underestimation of motion can be resolved by adjusting the free parameters and thresholds involved in various stages of the proposed procedure. This can be accomplished by testing the proposed procedure on a large database of video recordings of neonatal seizures and clinical events not due to seizures, which is currently in progress. This fine-tuning of the proposed procedure is expected to eliminate, or at least to reduce, the uncertainty associated with the validity of the resulting motion strength signals. In practice, the quantification of neonatal seizures can become more reliable by using temporal motion strength signals in conjunction with temporal motor activity signals, which quantify motor activity by relying on different but complementary computational procedures [5], [6].

The long-term goal of the research described in this paper is to integrate the computational procedures outlined in this paper into the development of a stand-alone automated system that could be used as a supplement in the neonatal intensive care unit to: 1) provide 24-hour a day noninvasive monitoring of infants at risk for seizures, and 2) facilitate the analysis and characterization of videotaped neonatal seizures by physicians during retrospective review. This goal will be accomplished by developing an intelligent system capable of recognizing focal clonic and myoclonic seizures and distinguishing them from videotaped clinical events characterized by increased motor activity of the infant's extremities.

5. References

[1] A. M. E. Bye, C. A. Cunningham, K. Y. Chee, and D. Flanagan, "Outcome of neonates with electrographically identified seizures, or at risk of seizures," *Pediatric Neurology,* vol. 16, pp. 225-231, 1990.

[2] G. M. Fenichel, *Neonatal Neurology,* 3rd ed., Churchill-Livingstone, New York, NY, 1990.

[3] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2002.

[4] K. R. Holden, E. D. Mellits, and J. M. Freeman, "Neonatal seizures. I. Correlation of prenatal and perinatal events with outcomes," *Pediatrics*, vol. 70, pp. 165-176, 1982.

[5] N. B. Karayiannis, "Advancing videometry through applications: Quantification of neonatal seizures from video recordings," *Proc. Fourteenth International Conference on Digital Signal Processing, Santorini*, Greece, July 1-3, 2002, pp. 11-21.

[6] N. B. Karayiannis, S. Srinivasan, R. Bhattacharya, 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," *IEEE Trans. on Medical Imaging,* vol. 20, no. 9, pp. 965-980, 2001.

[7] E. D. Mellitis, K. R. Holden, and J. M. Freeman, "Neonatal seizures. II. A multivariate analysis of factors associated with outcome," *Pediatrics*, vol. 70, pp. 177-185, 1982.

[8] F. Meyer and S. Beucher, "Morphological segmentation," *J. Visual Commun. Image Represent.*, vol. 1, pp. 21-46, 1990.

[9] E. M. Mizrahi, "Neonatal seizures," in *Pediatric and Adolescent Medicine*, S. Shinnar, N. Amir, D. Branski (Eds.), Basel, Karger, vol. 6, pp. 18-31, 1995.

[10] E. M. Mizrahi, "Acute and chronic effects of seizures in the developing brain: Lessons from clinical experience," *Epilepsia*, vol. 40 (Suppl 1), pp. 42-50, 1999.

[11] E. M. Mizrahi and P. Kellaway, "Characterization and classification of neonatal seizures," *Neurology*, vol. 37, pp. 1837-1844, 1987.

[12] G. M. Ronen and S. Penney, "The epidemiology of clinical neonatal seizures in Newfoundland, Canada: A five-year cohort," *Ann Neurol*, vol. 38, pp. 518-519, 1995

[13] A. R. Seay and P. F. Bray, "Significance of seizures in infants weighing less than 2500 grams," *Arch Neurol,* vol. 34, pp. 381-382, 1977.

[14] J. J. Volpe, *Neurology of the Newborn,* WB Suanders, Philadelphia, PA, 1995.