

# Extraction of Temporal Motor Activity Signals From Video Recordings of Neonatal Seizures By Feature Tracking Methods Based on Rigid Motion Models

Nicolaos B. Karayiannis and Abdul Sami

Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005

**Abstract**—This paper presents a new method for tracking features in video. This method estimates the displacement of a feature between two successive frames by minimizing an error function defined in terms of the feature intensities at these frames. Feature tracking relies on a rigid motion model, which allows for translation, rotation and uniform scaling of the feature tracked from one frame to the next. The proposed feature tracking method is used to extract motor activity signals from video recordings of neonatal seizures.

**Keywords**—Feature tracking, motor activity signal, rigid motion, rotation, scaling

## I. INTRODUCTION

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 [2], [6], [10]. Video recording is typically used with synchronized EEG and other polygraphic measures to analyze the characteristics of a seizure after its recording [1], [6], [7]. 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, and spasms), and reveal the precise sequence of motor components within a single seizure.

Motor activity due to neonatal seizures was quantified in a recent study by extracting temporal motion strength and motor activity signals [3], [4]. Temporal motor activity signals can be extracted from video by projecting to the horizontal and vertical axes an anatomical site located at a moving body part that may be affected by a seizure. The anatomical sites located at moving body parts were tracked in this study by employing the KLT algorithm, which determines the location of a feature in the next frame by estimating its displacement between two adjacent frames based on a motion model involving pure translation [5], [9]. The study outlined in this paper focused on the development of a feature tracker based on a rigid motion model that can be used to track features that may be translated, rotated, and uniformly scaled from one frame to the next.

---

This work was supported by the National Institute of Biomedical Imaging and Bioengineering under Grant 1 R01 EB00183-01.

## II. EXTRACTION OF TEMPORAL MOTOR ACTIVITY SIGNALS FROM VIDEO

Figure 1 illustrates the mechanism that can be used for generating temporal signals tracking the movements of different parts of the infant's body during focal clonic and myoclonic seizures [4]. Figure 1 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 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  $Y_{LH}$ , and  $X_{RH}$  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 videotaped 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.

## III. RIGID MOTION MODELS

Consider a frame sequence  $\{I(\mathbf{u}, t)\}$ , where  $u = [x \ y]^T$ ,  $\mathbf{a}^T$  denotes the transpose of a vector  $\mathbf{a}$ ,  $x$  and  $y$  are the coordinates of a pixel in the frame, and  $I(\mathbf{u}, t)$  represents the intensity of the pixel from frame  $t$  located at  $(x, y)$ . Let  $I(\mathbf{v}, t + \tau)$  be the intensity of a small region (i.e., a feature) at frame  $t + \tau$ , where  $\mathbf{v} = f(\mathbf{u})$  represents the new coordinates of the pixels within this region at frame  $t + \tau$ . The function  $f(\cdot)$  determines the model of motion employed for feature tracking.

Consider the linear model

$$\mathbf{v} = \mathbf{u} + \mathbf{A}\mathbf{u} + \mathbf{d}_u, \quad (1)$$

where  $\mathbf{A} \in \mathbb{R}^{2 \times 2}$  and  $\mathbf{d}_u = [d_x \ d_y]^T$ . The pure translation model is a special case of a linear motion model (1), which corresponds to  $\mathbf{A} = \mathbf{0}$ . The linear model (1) implements translation and a proper rotation if  $\mathbf{A}$  satisfies the constraints  $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ , and  $\det \mathbf{A} = 1$ . The linear model (1) implements translation, rotation, and uniform scaling if  $\mathbf{A}$

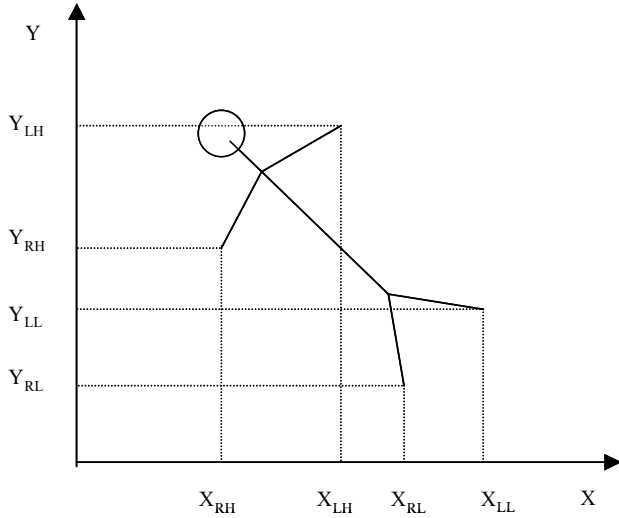


Figure 1: Extraction of temporal motor activity signals by projecting four selected anatomical sites to the horizontal and vertical axes.

satisfies the constraints  $\mathbf{A}^T \mathbf{A} = s^2 \mathbf{I}$  and  $\det \mathbf{A} = s^2$ , where  $s$  is the scaling factor. These constraints are satisfied by  $\mathbf{A} = s \mathbf{R}(\theta)$ , where

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos \theta & \mp \sin \theta \\ \pm \sin \theta & \cos \theta \end{bmatrix}. \quad (2)$$

If  $\mathbf{A} = s \mathbf{R}(\theta)$ , then the motion of the pixels within the feature tracked can be described by the vector  $\mathbf{z} = [\theta \ s \ d_x \ d_y]^T$ , which is formed in terms of the parameters of the motion model.

It is assumed that, under the linear motion model  $\mathbf{v} = \mathbf{u} + \mathbf{A}\mathbf{u} + \mathbf{d}_u$ , the intensities of the pixels within the feature tracked remain the same, that is,

$$I(\mathbf{z} + \delta \mathbf{z}, t + \tau) = I(\mathbf{z}, t). \quad (3)$$

The assumption is valid only for sufficiently high temporal sampling rates.

#### IV. FEATURE TRACKING

Tracking of a feature (i.e., block of pixels) throughout a sequence of frames requires the development of a procedure for estimating the unknown vector  $\delta \mathbf{z}$  between two successive frames in terms of the pixel intensities in these frames. This can be accomplished by minimizing the error [5], [9]

$$\varepsilon = \frac{1}{2} \sum_W [I(\mathbf{z} + \delta \mathbf{z}, t + \tau) - I(\mathbf{z}, t)]^2, \quad (4)$$

where  $W$  is a window located at the center pixel of the feature tracked. The minimization of  $\varepsilon$  can be made analytically tractable by approximating  $I(\mathbf{z} + \delta \mathbf{z}, t + \tau)$  using a first-order Taylor expansion about  $\mathbf{z}$  as

$$I(\mathbf{z} + \delta \mathbf{z}, t + \tau) = I(\mathbf{z}, t + \tau) + \mathbf{g}_z^T \delta \mathbf{z}, \quad (5)$$

where  $\mathbf{g}_z \doteq \nabla_{\mathbf{z}} I(\mathbf{z}, t + \tau)$  denotes the gradient of  $I(\cdot)$  with respect to  $\mathbf{z}$ , defined as

$$\mathbf{g}_z = \left[ \frac{\partial I}{\partial \theta} \quad \frac{\partial I}{\partial s} \quad \frac{\partial I}{\partial d_x} \quad \frac{\partial I}{\partial d_y} \right]^T. \quad (6)$$

Using this approximation, the error defined in (4) becomes

$$\varepsilon = \frac{1}{2} \sum_W [I(\mathbf{z}, t + \tau) - I(\mathbf{z}, t) + \mathbf{g}_z^T \delta \mathbf{z}]^2. \quad (7)$$

The unknown vector  $\delta \mathbf{z}$  can be obtained by solving the equation

$$\nabla_{\delta \mathbf{z}} \varepsilon = \sum_W \mathbf{g}_z [I(\mathbf{z}, t + \tau) - I(\mathbf{z}, t) + \mathbf{g}_z^T \delta \mathbf{z}] = \mathbf{0}. \quad (8)$$

The equation (8) can also be written as

$$\mathbf{G} \delta \mathbf{z} = \mathbf{e}, \quad (9)$$

where

$$\mathbf{G} = \sum_W \mathbf{g}_z \mathbf{g}_z^T, \quad (10)$$

and

$$\mathbf{e} = \sum_W \mathbf{g}_z [I(\mathbf{z}, t) - I(\mathbf{z}, t + \tau)]. \quad (11)$$

The estimate  $\delta \mathbf{z}$  obtained by solving (9) may not be particularly accurate. An alternative approach is to minimize the error in (7) by using an iterative optimization procedure, such as the Newton-Raphson method. In such a case, the new estimate  $\delta \mathbf{z}^{new}$  of the unknown vector  $\delta \mathbf{z}$  is obtained in terms of the current estimate  $\delta \mathbf{z}^{old}$  as

$$\delta \mathbf{z}^{new} = \delta \mathbf{z}^{old} - \mathbf{H}^{-1} \nabla_{\delta \mathbf{z}} \varepsilon, \quad (12)$$

where  $\nabla_{\delta \mathbf{z}} \varepsilon = -(\mathbf{e} - \mathbf{G} \delta \mathbf{z})$  is the gradient of  $\varepsilon$  with respect to  $\delta \mathbf{z}$ , and  $\mathbf{H}$  is the Hessian matrix. For the error function defined in (7), the Hessian matrix can be obtained as

$$\mathbf{H} = \mathbf{G} = \sum_W \mathbf{g}_z \mathbf{g}_z^T. \quad (13)$$

For  $\nabla_{\delta \mathbf{z}} \varepsilon = -(\mathbf{e} - \mathbf{G} \delta \mathbf{z})$  and  $\mathbf{H}^{-1} = \mathbf{G}^{-1}$ , the update equation (12) becomes

$$\delta \mathbf{z}^{new} = \mathbf{G}^{-1} \mathbf{e}. \quad (14)$$

In this particular case, each iteration of the Newton-Raphson method is equivalent to solving the equation (9).

#### IV. FEATURE TRACKING BASED ON RIGID MOTION MODELS

Consider the rigid motion model  $\mathbf{v} = \mathbf{u} + s \mathbf{R}(\theta) \mathbf{u} + \mathbf{d}_u$ . Since  $\mathbf{v} = \phi(\mathbf{u}, \mathbf{z})$ ,

$$\nabla_{\mathbf{z}} I(\mathbf{z}, t + \tau) = \nabla_{\mathbf{z}}(\mathbf{v})^T \nabla_{\mathbf{v}} I(\mathbf{v}, t + \tau), \quad (15)$$

where

$$\nabla_{\mathbf{z}}(\mathbf{v}) = \left[ \frac{\partial \mathbf{v}}{\partial \theta} \quad \frac{\partial \mathbf{v}}{\partial s} \quad \frac{\partial \mathbf{v}}{\partial d_x} \quad \frac{\partial \mathbf{v}}{\partial d_y} \right]. \quad (16)$$

The gradient  $\nabla_{\mathbf{v}} I(\mathbf{v}, t + \tau)$  can be computed in terms of the gradient  $\mathbf{g}_{\mathbf{u}} \doteq \nabla_{\mathbf{u}} I(\mathbf{u}, t)$ , defined as

$$\mathbf{g}_{\mathbf{u}} = \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix}^T. \quad (17)$$

This can be accomplished by using (3), which can also be written as

$$I(\mathbf{v}, t + \tau) = I(\mathbf{u}, t). \quad (18)$$

Since  $\mathbf{v} = \phi(\mathbf{u}, \mathbf{z})$ , taking the gradient with respect to  $\mathbf{u}$  of both sides of (18) gives

$$\nabla_{\mathbf{u}}(\mathbf{v})^T \nabla_{\mathbf{v}} I(\mathbf{v}, t + \tau) = \nabla_{\mathbf{u}} I(\mathbf{u}, t), \quad (19)$$

where

$$\nabla_{\mathbf{u}}(\mathbf{v}) = \begin{bmatrix} \frac{\partial \mathbf{v}}{\partial x} & \frac{\partial \mathbf{v}}{\partial y} \end{bmatrix}. \quad (20)$$

Since  $\mathbf{v} = \mathbf{u} + s \mathbf{R}(\theta) \mathbf{u} + \mathbf{d}_{\mathbf{u}}$ ,  $\nabla_{\mathbf{u}}(\mathbf{v}) = \mathbf{I} + s \mathbf{R}(\theta)$ . Thus, (19) gives

$$\nabla_{\mathbf{v}} I(\mathbf{v}, t + \tau) = [(\mathbf{I} + s \mathbf{R}(\theta))^{-1}]^T \mathbf{g}_{\mathbf{u}}. \quad (21)$$

The gradient  $\mathbf{g}_{\mathbf{z}} = \nabla_{\mathbf{z}} I(\mathbf{z}, t + \tau)$  can be obtained by combining (21) and (15) as

$$\mathbf{g}_{\mathbf{z}} = [(\mathbf{I} + s \mathbf{R}(\theta))^{-1} \nabla_{\mathbf{z}}(\mathbf{v})]^T \mathbf{g}_{\mathbf{u}}. \quad (22)$$

For  $\mathbf{v} = \mathbf{u} + s \mathbf{R}(\theta) \mathbf{u} + \mathbf{d}_{\mathbf{u}}$ , it can be shown that

$$[\mathbf{I} + s \mathbf{R}(\theta)]^{-1} \nabla_{\mathbf{z}}(\mathbf{v}) = \frac{1}{s^2 + 2s \cos \theta + 1} [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3 \quad \mathbf{h}_4], \quad (23)$$

where

$$\mathbf{h}_1 = \begin{bmatrix} -sx \sin \theta \mp sy(\cos \theta + s) \\ \pm sx(\cos \theta + s) - sy \sin \theta \end{bmatrix}, \quad (24)$$

$$\mathbf{h}_2 = \begin{bmatrix} x(\cos \theta + s) \mp y \sin \theta \\ \pm x \sin \theta + y(\cos \theta + s) \end{bmatrix}, \quad (25)$$

$$\mathbf{h}_3 = \begin{bmatrix} 1 + s \cos \theta \\ \mp s \sin \theta \end{bmatrix}, \quad (26)$$

$$\mathbf{h}_4 = \begin{bmatrix} \pm s \sin \theta \\ 1 + s \cos \theta \end{bmatrix}. \quad (27)$$

## V. EXPERIMENTAL RESULTS

Figures 2 and 3 show the motor activity signals extracted from the video recordings of neonatal seizures by utilizing the feature tracking methods based on pure translation and rigid motion models. The locations of the moving body parts during the clinical event are shown in representative frames of each video recording. The values of the signals corresponding to the frames shown at the top of each figure are indicated by dots, while the features tracked in each video recording are indicated by a box.

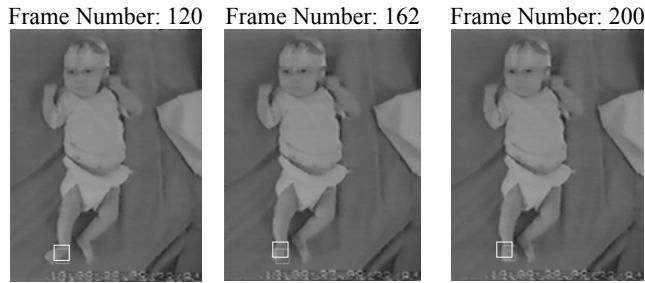
Figure 2 shows the temporal motor activity signals produced for a myoclonic seizure affecting the infant's right foot by a feature tracking method relying on a pure

translation and the feature tracking method developed in this paper based on a rigid motion model. Both methods identified significant motor activity along the horizontal direction around frame 160. Frame-by-frame inspection of the video recording indicated that there was no substantial motor activity outside a short interval in the neighborhood of frame 160. Thus, the fluctuations of the motor activity signals outside this interval are indeed artifacts of the feature tracking methods. Nevertheless, the appearance of such fluctuations in the motor activity signals does not actually affect the quantification of substantial motor activity that may be due to seizures. In fact, the signals quantifying motor activity along the horizontal direction are consistent with the rapid and "jerky" motion due to the myoclonic seizure.

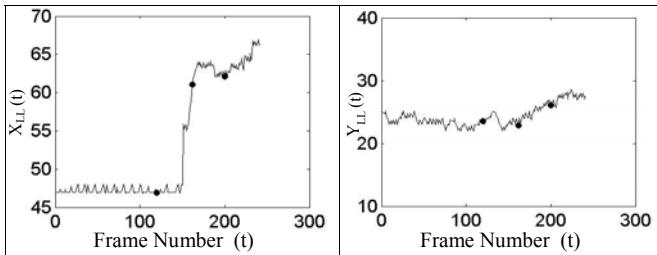
Figure 3 shows the temporal motor activity signals produced by the two aforementioned feature tracking methods tested on the video recording of a focal clonic seizure affecting the infant's right leg. Both feature tracking methods identified substantial motor activity along the vertical direction. However, the two methods produced different motor activity signals especially after frame 150. Frame-by-frame inspection of the video recording indicated that the motor activity signal produced by the proposed feature tracking method constitutes a better representation of the actual motor activity after frame 150. Regardless of their differences, the motor activity signals shown in Figure 3 captured the rhythmicity that is the signature characteristic of focal clonic seizures.

## VI. CONCLUSIONS

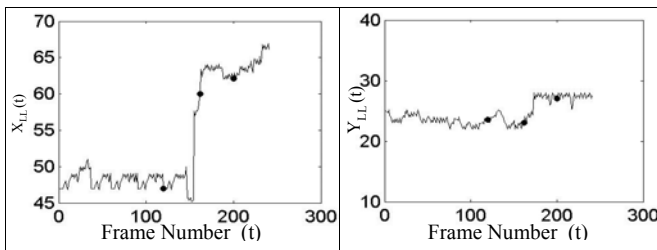
This paper introduced a new feature tracking method for video. This method relies on a rigid motion model that can track features subjected to translation, rotation, and uniform scaling between two successive frames. The proposed feature tracking method is the essential generalization of feature tracking methods relying on a pure translation model. In terms of its computational requirements, the proposed feature tracking method is a compromise between feature tracking methods that rely on pure translation motion models and those relying on affine motion models [8]. On the other hand, the proposed method would not be capable of tracking features that may be deformed from one frame to the next. Nevertheless, it seems that this shortcoming is not particularly significant for the extraction of temporal motor activity signals from video recordings of neonatal seizures. The reason is that the features tracked in this application typically occupy a relatively small region located at the infant's moving body part. This implies that this application may not require sophisticated motion models to quantify the motion of the features tracked between successive frames. This argument is also supported by the outcome of the experiments reported in this paper, which revealed that there are no significant differences between the motor activity signals produced for the video recordings of two neonatal seizures by the feature tracking methods relying on



(a)



(b)



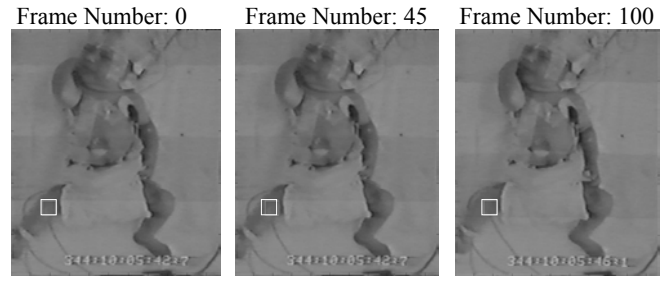
(c)

Figure 2: (a) Selected frames of a video recording of a myoclonic seizure affecting the infant's right leg, (b) motor activity signals produced by the feature tracking method relying on a pure translation model, (c) motor activity signals produced by a feature tracking method relying on a rigid motion model.

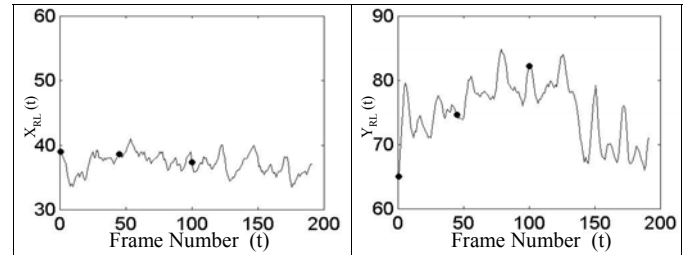
pure translation and rigid motion models. The comparison of these feature tracking methods is the focus of a major study currently under way, which involves their testing on a large database of video recordings of neonatal seizures and other normal and abnormal infant behaviors not due to seizures.

#### REFERENCES

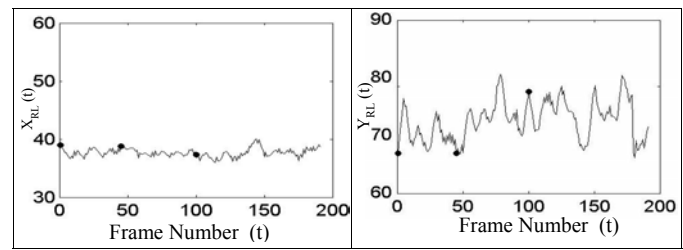
- [1] A. M. E. Bye, P. Lamont, L. Healy, "Commencement of a pediatric EEG-video telemetry service," *Clin Exp Neurol*; vol. 27, pp. 83-88, 1990.
- [2] G. M. Fenichel, *Neonatal Neurology*, 3rd ed. New York, NY: Churchill-Livingstone, 1990.
- [3] N. B. Karayiannis, "Advancing videometry through applications: Quantification of neonatal seizures from video recordings," *Proceedings of Fourteenth International Conference on Digital Signal Processing*, Santorini, Greece, July 1-3, 2002, pp. 11-21.
- [4] 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 Transactions on Medical Imaging*, vol. 20, no. 9, pp. 965-980, 2001.
- [5] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereoscopic vision," *Proceedings of International Joint Conference on Artificial Intelligence*, Vancouver, Canada, August 1981, pp. 674-679.
- [6] 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.
- [7] E. M. Mizrahi and P. Kellaway, "Characterization and classification of neonatal seizures," *Neurology*, vol. 37, pp. 1837-1844, 1987.
- [8] J. Shi and C. Tomasi, "Good features to track," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WA, June 1994, pp. 593-600.
- [9] C. Tomasi and T. Kanade, "Detection and tracking of point features," *Carnegie Mellon University Technical Report CMU-CS-91-132*, April 1991.
- [10] J. J. Volpe. *Neurology of the Newborn*. Philadelphia: WB Saunders, 1995.



(a)



(b)



(c)

Figure 3: (a) Selected frames of a video recording of a focal clonic seizure affecting the infant's right leg, (b) motor activity signals produced by the feature tracking method relying on a pure translation model, (c) motor activity signals produced by a feature tracking method relying on a rigid motion model.