

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

Nicolaos B. Karayiannis and Yaohua Xiong

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. The minimization problem is made analytically tractable by approximating the error function using a second-order Taylor expansion. The displacement between two successive frames is computed in an iterative fashion using gradient descent. The improved reliability of the proposed method is illustrated by its application in the extraction of temporal motor activity signals from video recordings of neonatal seizures.

Keywords—Feature tracking, motor activity signal, translation motion model

I. INTRODUCTION

Feature tracking is an essential operation in many applications involving video surveillance. The problem is to track a feature (i.e., a block of pixels) throughout a sequence of frames. A typical application involving feature tracking is the extraction of temporal motor activity signals from video recording of neonatal seizures [2], [3]. These signals can be obtained by projecting to the horizontal and vertical axes an anatomical site located at a moving body part that may be affected by a seizure. Feature tracking can be performed by the KLT algorithm, which was developed to track “good features” throughout a frame sequence [4], [8].

The KLT algorithm was utilized recently to extract motor activity signals from video recordings of neonatal seizures [2], [3]. Although the KLT algorithm was generally successful, in some cases the algorithm lost features that were located at moving body parts tracked throughout the frame sequence. A feature was lost when the error function computed between two successive frames assumed a value higher than a certain threshold. The susceptibility of the KLT algorithm to “lost features” was dealt with by tracking a sufficiently large number of features within a predetermined radius from the selected anatomical site.

This paper proposes a new procedure for estimating the location of a feature in the next frame of the video recording. The displacement of the feature between two

successive frames is estimated by minimizing an error function defined in terms of the intensity functions at these frames. In the proposed procedure, the error function is approximated by using a second-order Taylor expansion for the intensity function at the next frame. The proposed feature tracking method is used to extract motor activity signals from video recordings of neonatal seizures.

II. EXTRACTION OF MOTOR ACTIVITY SIGNALS FROM VIDEO

Motor activity signals can be extracted by projecting the location of selected anatomical sites to the horizontal and vertical axes. As the seizure progresses in time, these projections will produce temporal signals recording motor activity of the body parts of interest.

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. Focal clonic and myoclonic seizures are manifested as repetitive and rapid movements of the infants’ extremities, respectively [1], [5], [6], [10]. 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. FEATURE TRACKING

Consider a frame sequence $\{I(\mathbf{u}, t)\}$, where $\mathbf{u} = [x \ y]^T$, \mathbf{a}^T denotes the transpose of a vector \mathbf{a} , x and y are the coordinates of a pixel in the frame. It is assumed that the intensities of small frame regions are displaced but their intensities remain the same, that is,

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

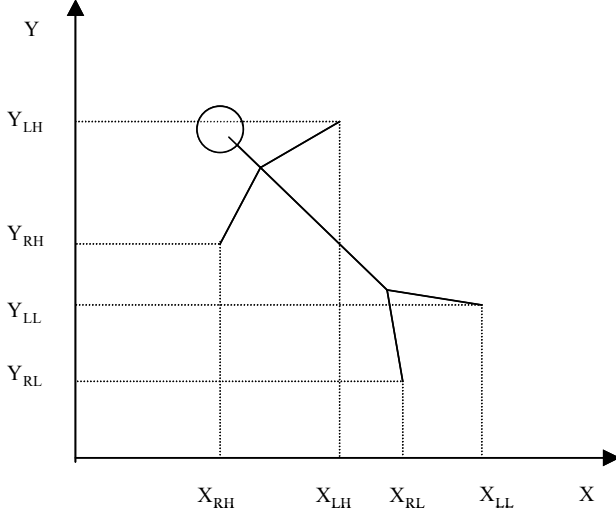


Figure 1: Extraction of temporal motor activity signals by projecting four selected anatomical sites to the horizontal and vertical axes.

$$I(\mathbf{u} + \mathbf{d}, t + \tau) = I(\mathbf{u}, t), \quad (1)$$

where $\mathbf{d} = [d_x \ d_y]^T$ is the displacement vector. The condition in (1) is valid under the assumption that motion can be approximated by pure translation. This assumption is valid only for sufficiently high temporal sampling rates. Tracking of a feature (i.e., a block of pixels) requires the development of a procedure for estimating the displacement \mathbf{d} of the feature between two successive frames from the pixel intensities in these frames. The procedure employed by the KLT algorithm for feature tracking was developed by minimizing the error [4], [8]

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

where W is a window centered at the pixel whose displacement \mathbf{d} is estimated. The minimization of ε was made analytically tractable by approximating $I(\mathbf{u} + \mathbf{d}, t + \tau)$ using a first-order Taylor expansion as [4], [8]

$$I(\mathbf{u} + \mathbf{d}, t + \tau) = I(\mathbf{u}, t + \tau) + \mathbf{g}^T \mathbf{d}, \quad (3)$$

where $\mathbf{g} = [I_x \ I_y]^T$ and $I_x = \partial I / \partial x$, $I_y = \partial I / \partial y$. Using this approximation, the error defined in (2) becomes

$$\varepsilon = \frac{1}{2} \sum_W [I(\mathbf{u}, t + \tau) - I(\mathbf{u}, t) + \mathbf{g}^T \mathbf{d}]^2. \quad (4)$$

The displacement vector \mathbf{d} can be obtained in terms of the gradient $\nabla_{\mathbf{d}} \varepsilon$ of ε with respect to \mathbf{d} by solving the equation

$$\nabla_{\mathbf{d}} \varepsilon = \sum_W \mathbf{g} [I(\mathbf{u}, t + \tau) - I(\mathbf{u}, t) + \mathbf{g}^T \mathbf{d}] = \mathbf{0}. \quad (5)$$

The equation (5) can also be written as

$$\mathbf{G} \mathbf{d} = \mathbf{e}, \quad (6)$$

where

$$\mathbf{G} = \sum_W \mathbf{g} \mathbf{g}^T, \quad (7)$$

and

$$\mathbf{e} = \sum_W \mathbf{g} [I(\mathbf{u}, t) - I(\mathbf{u}, t + \tau)]. \quad (8)$$

The displacement \mathbf{d} can be updated several times in order to reduce the error ε . A feature is typically rejected if the error ε resulting after several updates is higher than a certain threshold.

IV. A NEW FEATURE TRACKING METHOD

The approach employed by the KLT algorithm for estimating the displacement \mathbf{d} can be improved by minimizing the same error ε , defined in (2), but using a better approximation for $I(\mathbf{u} + \mathbf{d}, t + \tau)$. More specifically, $I(\mathbf{u} + \mathbf{d}, t + \tau)$ can be approximated by using a second-order Taylor expansion about \mathbf{u} as

$$I(\mathbf{u} + \mathbf{d}, t + \tau) = I(\mathbf{u}, t + \tau) + \mathbf{g}^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \mathbf{H} \mathbf{d}, \quad (9)$$

where \mathbf{H} is defined in terms of $I_{xx} = \partial^2 I / \partial x^2$, $I_{xy} = \partial^2 I / \partial x \partial y$, $I_{yx} = \partial^2 I / \partial y \partial x$, and $I_{yy} = \partial^2 I / \partial y^2$ as

$$\mathbf{H} = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix}. \quad (10)$$

Using this approximation, the error ε defined in (2) takes the form

$$\varepsilon = \frac{1}{2} \sum_W [I(\mathbf{u}, t + \tau) - I(\mathbf{u}, t) + \mathbf{g}^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \mathbf{H} \mathbf{d}]^2. \quad (11)$$

The gradient $\nabla_{\mathbf{d}} \varepsilon$ of ε with respect to \mathbf{d} is

$$\nabla_{\mathbf{d}} \varepsilon = \sum_W [I(\mathbf{u}, t + \tau) - I(\mathbf{u}, t) + \mathbf{g}^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T \mathbf{H} \mathbf{d}] (\mathbf{g} + \mathbf{H} \mathbf{d}). \quad (12)$$

In this case, \mathbf{d} cannot be obtained directly by solving the equation $\nabla_{\mathbf{d}} \varepsilon = \mathbf{0}$. Instead, \mathbf{d} can be obtained in an iterative fashion by minimizing the error ε using gradient descent. According to such an approach, the new value \mathbf{d}^{new} of the displacement vector \mathbf{d} after each iteration can be obtained in terms of the value \mathbf{d}^{old} of \mathbf{d} produced by the previous iteration as

$$\mathbf{d}^{\text{new}} = \mathbf{d}^{\text{old}} - \alpha \nabla_{\mathbf{d}} \varepsilon \Big|_{\mathbf{d}=\mathbf{d}^{\text{old}}}, \quad (13)$$

where $\nabla_{\mathbf{d}} \varepsilon$ is the gradient in (12) and α is the step size.

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 method employed by the KLT algorithm, a feature tracking method employing a first-order Taylor expansion and gradient descent for minimization and a feature tracking method employing a second-order Taylor expansion and gradient descent for minimization. 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 and 3 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). This movement was captured by both feature tracking methods that relied on gradient descent for minimization. The same two methods captured the movement of the left leg toward the top of the frame, as indicated by the temporal motor activity signals obtained as the projection of the feature tracked to the vertical axis. However, the motor activity occurring between frames 10 and 16 is not shown in Figure 2(b), which was produced by using the feature tracking method employed by the KLT algorithm. The reason is that the feature located at the infant's left leg was lost by this feature tracking method after frame 13. In fact, the feature tracked by this method after frame 13 was located at the bed and remained almost fixed throughout the sequence; this explains the flat motor activity signals shown in Figure 2(b). The infant's left leg remains at an almost fixed position between frames 50 and 150. In this time interval, the temporal motor activity signals produced by the two feature tracking methods based on gradient descent are almost flat. 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.

Figure 3 shows the temporal motor activity signals produced by the feature tracking methods tested in the experiments for a focal clonic seizure affecting the infant's right leg. Figure 3 indicates that the temporal signals produced by all three feature tracking methods capture and quantify the rhythmicity that is the signature characteristic of focal clonic seizures. In fact, all these methods were successful in tracking the feature located at the infant's right leg. However, the three feature tracking methods produced different measurements of motor activity as indicated by the different amplitudes of the temporal signals representing motor activity along the vertical direction. In fact, the feature tracking method employed by the KLT algorithm produced the motor activity signals of the smallest amplitude. On the other hand, the motor activity signals with the largest amplitude were those produced by the feature tracking method relying on a second-order Taylor expansion for approximation and on gradient descent for minimization.

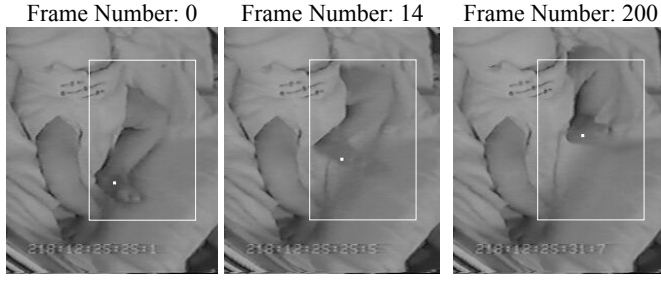
Frame-by-frame visual inspection of the video recording indicated that the feature tracked moved considerably along the vertical direction between frames 60 and 160. This is consistent with the motor activity signals produced by the feature tracking methods based on gradient descent and shown in Figures 3(c) and 3(d).

VI. CONCLUSIONS

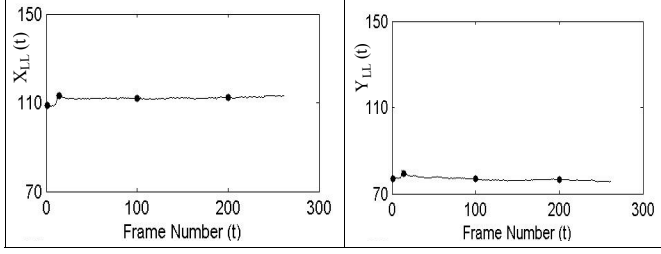
This paper introduced a new method for tracking features in video. The proposed method relies on a pure translation motion model and estimates the displacement of a feature between two successive frames by minimizing an error function defined in terms of the feature intensities at these frames. The minimization problem was made analytically tractable by approximating the error function using a second-order Taylor expansion. The displacement was computed in an iterative fashion using gradient descent. The proposed method was evaluated and compared with other feature tracking methods on the extraction of temporal motor activity signals from video recordings of neonatal seizures. The experiments indicated that the proposed method outperformed considerably two alternative feature tracking methods that rely on a first-order Taylor expansion to approximate the error function. This experimental outcome indicates that the superiority of the proposed method can only be attributed to the approximation of the error function by means of a second-order Taylor expansion. An interesting problem for future research is to combine the proposed treatment of the minimization problem with a motion model that is more sophisticated than the pure translation model considered in this paper. This can be accomplished by employing rigid or deformable motion models [7], [9].

REFERENCES

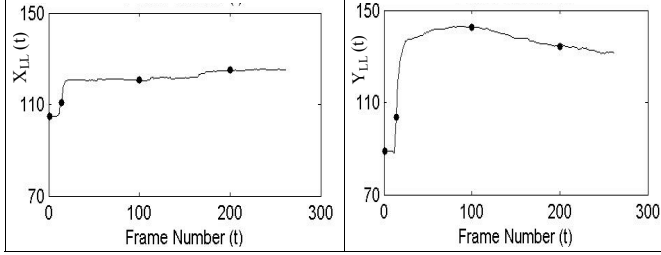
- [1] G. M. Fenichel, *Neonatal Neurology*, 3rd ed. New York, NY: Churchill-Livingstone, 1990.
- [2] N. B. Karayiannis, "Advancing videometry through applications: Quantification of neonatal seizures from video recordings," *Proceedings of Fourteenth International Conference on Digital Signal Processing*, Santorini, Greece, July 1-3, 2002, pp. 11-21.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] E. M. Mizrahi and P. Kellaway, "Characterization and classification of neonatal seizures," *Neurology*, vol. 37, pp. 1837-1844, 1987.
- [7] 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.



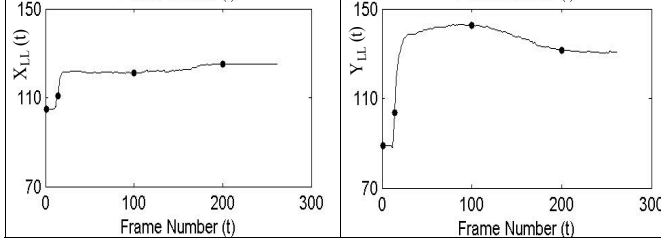
(a)



(b)

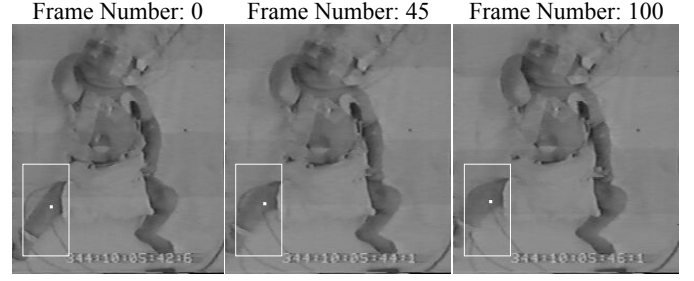


(c)

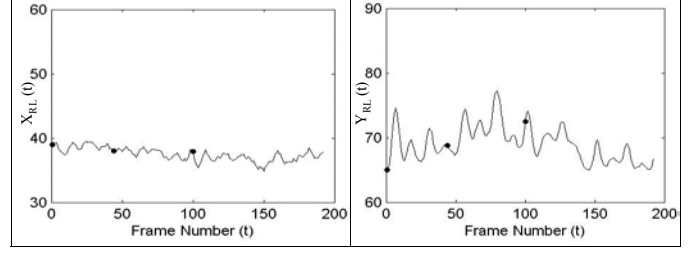


(d)

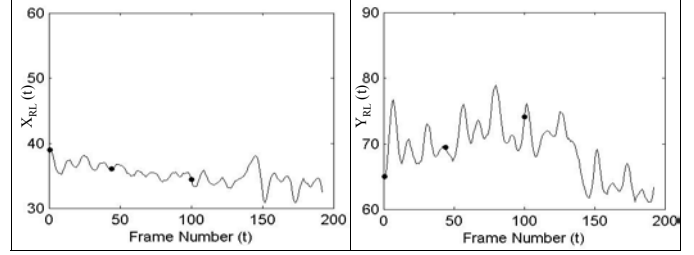
Figure 2: (a) Selected frames of a video recording of a myoclonic seizure affecting the infant's left leg, (b) motor activity signals produced by the feature tracking method employed by the KLT algorithm, (c) motor activity signals produced by a feature tracking method employing a first-order Taylor expansion and gradient descent for minimization, and (d) motor activity signals produced by a feature tracking method employing a second-order Taylor expansion and gradient descent for minimization.



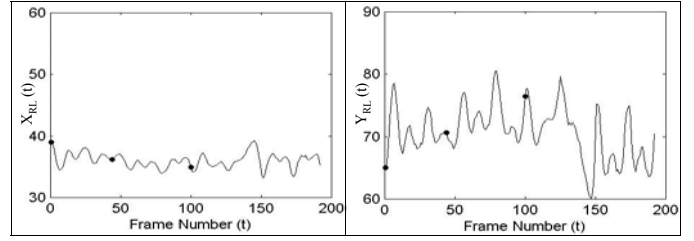
(a)



(b)



(c)



(d)

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 employed by the KLT algorithm, (c) motor activity signals produced by a feature tracking method employing a first-order Taylor expansion and gradient descent for minimization, and (d) motor activity signals produced by a feature tracking method employing a second-order Taylor expansion and gradient descent for minimization.

- [8] C. Tomasi and T. Kanade, "Detection and tracking of point features," *Carnegie Mellon University Technical Report CMU-CS-91-132*, April 1991.
- [9] T. Tommasini, A. Fusiello, E. Trucco, and V. Roberto, "Making good features track better," *Proceedings of IEEE Conference on Computer*

- Vision and Pattern Recognition*, Santa Barbara, CA, June 1998, pp. 178-183.
- [10] J. J. Volpe. *Neurology of the Newborn*. Philadelphia: WB Saunders, 1995.