Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Image and Vision Computing 24 (2006) 27–40
image AND vision COMPUTING
www.elsevier.com/locate/imavis

An improved procedure for the extraction of temporal motion strength signals from video recordings of neonatal seizures

Nicolaos B. Karayiannis *, Guozhi Tao

Department of Electrical and Computer Engineering, University of Houston, N308 Engineering Building 1, Houston, TX 77204-4005, USA

Received 22 December 2004; received in revised form 8 September 2005; accepted 15 September 2005

Abstract

This paper presents a procedure developed to extract quantitative motion information from video recordings of neonatal seizures in the form of temporal motion strength signals. Temporal motion strength signals are obtained from a sequence of video frames by measuring the displacement areas of the infants' moving body part(s) from frame to frame of the video sequence. The proposed motion segmentation procedure relies on the application of non-linear filtering, vector clustering, and morphological filtering to the differences between adjacent frames. The experiments indicate that temporal motion strength signals constitute a satisfactory representation of videotaped clinical events and may be used for automated seizure recognition.

Keywords: Clustering; Median filtering; Morphological filtering; Motion segmentation; Motion strength signal; Neonatal seizure; Video recording

1. Introduction

Seizures occur in approximately 2–5/1000 live births, depending upon studied populations and methodology [1–6]. In fact, the incidence of seizures in infants weighing less than 1500 g is 57.5/1000 live births compared to 3.5/1000 live births for all birthweights [4]. Similarly, Scher et al. [7] reported that seizures occurred in approximately 4% of premature infants less than 30 weeks conceptional age, although some have reported the incidence in this population to be as high as 20%. These studies indicate that seizure occurrence represents the most frequent clinical sign of central nervous system dysfunction in the newborn [8,9]. 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.

Despite the importance of seizure recognition, most neonatal intensive care units and nurseries have limited resources for seizure identification. Neonatal seizures are often brief and may not be recognized since nurses and physicians cannot provide continuous surveillance of all infants at risk for clinical seizures. 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. Early attempts to characterize neonatal seizures involved primarily bedside observation and brief EEG recordings. Portable EEG/video/-polygraphic monitoring techniques have allowed investigators to assess and characterize neonatal seizures at the bedside and have permitted retrospective review. However, these techniques are relatively expensive, are generally used for only a few hours of monitoring, and are not routinely available in many centers [10–17].

The long-term goal of the study outlined in this paper is 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-h a day non-invasive monitoring of infants at risk for seizures, and (2) facilitate the analysis and characterization of videotaped neonatal seizures by physicians during retrospective review. The development of such a system requires automated procedures for extracting quantitative motion information from video recordings of infants monitored for seizures. The study described in this paper involved short video recordings of neonatal seizures of the myoclonic and focal clonic types, which affect the infants' extremities [14], and random movements of the infants' extremities not associated with seizures.

Motion in video recordings of infants monitored for seizures can be quantified by extracting temporal motion activity and

* Corresponding author. Tel.: +1 713 743 4436; fax: +1 713 743 4444.
E-mail address: karayiannis@uh.edu (N.B. Karayiannis).

0262-8856/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2005.09.009

motion strength signals [18,19]. Strength of motion mainly refers to the speed of the infants' moving body part(s), which can be indirectly measured by estimating the displacement areas of the infants' moving body part(s) from frame to frame of a video sequence. As the seizure progresses in time, the displacement area measurements A produce temporal motion strength signals $A(t)$. The extraction of motion strength signals requires an automated procedure capable of segmenting the displacement of the infants' moving body part(s) at each frame of the sequence. Segmentation of motion in an image sequence is one of the most challenging problems in computer vision. Segmentation based on motion information is unlikely to achieve an accurate result without the help of spatial information. In fact, many researchers proposed spatio-temporal segmentation procedures where spatial segmentation played an important role in the overall process [20–24].

Thompson [20] was the first to use a similarity constraint based on contrast and motion to achieve motion segmentation. This algorithm relies on a region-merging procedure, which is realized in two phases: In the first phase, adjacent regions with the same velocity label are always merged. Unlabeled regions and regions with different labels can be merged but only if they satisfy similarity conditions based on contrast information. In the second phase, velocity estimates are well determined so regions with different labels are not allowed to merge. The merging operation is applied only on the regions that satisfy the similarity conditions based on contrast information but have the same velocity label.

Dufaux et al. [21] proposed a new algorithm for the representation of a scene in terms of a moving object. Spatial pre-filtering is applied in the beginning to facilitate segmentation of the frames by producing constant luminance regions. The morphological OPENING and CLOSING operators are adopted for this task. Following the pre-processing stage, the algorithm begins with static segmentation and splits or merges regions based on motion information. Static segmentation is performed by applying the k-means clustering algorithm on the luminance values. For each of the resulting static regions, affine motion parameters are computed. Motion estimation is applied on a region characterized by coherent motion. Regions that are not well compensated are further split. Afterward, regions with similar motion are merged in the motion parameter space by applying clustering. Clustering in the motion parameter space results in regions characterized by coherent motion that can be identified as the moving object.

Choi et al. [22] presented a morphological spatio-temporal segmentation algorithm that incorporates luminance and motion information simultaneously and uses morphological tools such as morphological filters and the watershed algorithm. The procedure consists of three steps: joint marker extraction, boundary decision, and motion-based region fusion. The algorithm is implemented as follows: The frames are operated by the morphological OPENING and CLOSING filters to facilitate the segmentation process. The intensity markers are identified by labeling flat regions that have size larger than a given threshold. Then the motion markers are extracted to split the intensity marker for which the affine motion parameters are not accurate enough. The extraction of motion and intensity markers is followed by the boundary decision stage, which is used to deal with pixels not yet assigned to any region. The watershed algorithm is used as a region-growing tool based on the joint similarity measure. The joint similarity measure used for spatio-temporal segmentation is the weighted sum of the intensity difference and the motion difference. Finally, redundant regions are eliminated using motion-based region fusion. Regions created at stage two are merged together based on the consistency of an affine transformation.

Fusion of motion segmentation and image spatial segmentation can solve the over-segmentation problem since it relies on illumination segmentation to determine the object boundary. However, despite its advantages, fusion of motion segmentation and spatial segmentation is likely to trigger repeated split-and-merge procedures; this could increase the computational requirements of the resulting algorithms. If the motion boundary is not consistent with the intensity values of the object boundary, which is possible during complex movements, the results of spatial segmentation could mislead the motion segmentation algorithm. This may be the case in video recordings of neonatal seizures if only some small part of the infant's limb is moving,

The extraction of quantitative motion information from video recordings of neonatal seizures was initially attempted by using a motion segmentation procedure based on subband decomposition of video, which was followed by non-linear filtering and scalar clustering [18]. This procedure was tested on a few video recordings during an early study and produced motion strength signals consistent with the videotaped clinical event. However, additional testing revealed that the original procedure is sensitive to noise.

This paper presents an improved motion segmentation procedure developed for the extraction of motion strength signals. Motion segmentation was attempted in this study by combining the most promising among the tools and methodologies mentioned above into a motion segmentation procedure especially tailored to quantify video recordings of neonatal seizures. The proposed procedure was designed to overcome some of the deficiencies of the original procedure developed for the analysis of video recordings of neonatal seizures [18]. An additional constraint imposed on the development of the proposed motion segmentation procedure was that of low computational and storage requirements. This constraint was deemed necessary to ensure that the system under development would be capable of analyzing videotaped neonatal seizures in real time.

2. Methods

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 [18,19]. The extraction from video recordings of visual information that is relevant only to the seizure can be accomplished by the procedure outlined in this section.