Image and Vision Computing 26 (2008) 1621–1635

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

# Extraction and temporal segmentation of multiple motion trajectories in human motion

Junghye Min [a,*,1], Rangachar Kasturi [b], Octavia Camps [c,d]

[a] Digital Media R&D Center, Samsung Electronics, Suwon, Korea
[b] Department of Computer Science and Engineering,The University of South Florida,Tampa, FL, United States
[c] Department of Electrical Engineering, The Pennsylvania State University,University Park, PA, United States
[d] Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, United States

ARTICLE INFO

*Article history:*
Received 8 October 2005
Received in revised form 2 February 2008
Accepted 4 March 2008

*Keywords:*
Activity recognition
Motion trajectories
Motion tracking
Motion segmentation
Motion detection
Temporal segmentation

ABSTRACT

A new method for extraction and temporal segmentation of multiple motion trajectories in human motion is presented. The proposed method extracts motion trajectories generated by body parts without any initialization or any assumption on color distribution. Motion trajectories are very compact and representative features for activity recognition. Tracking human body parts (hands and feet) is inherently difficult because the body parts which generate most of the motion trajectories are relatively small compared to the human body. This problem is overcome by using a new motion segmentation method: at every frame, candidate motion locations are detected and set as significant motion points (SMPs). The motion trajectories are obtained by combining these SMPs and the color–optical flow based tracker results. These motion trajectories are inturn used as features for temporal segmentation of specific activities from continuous video sequences. The proposed approach is tested on actual ballet step sequences. Experimental results show that the proposed method can successfully extract and temporally segment multiple motion trajectories from human motion.



## 1. Introduction

As more and more video data are available everywhere, there is growing interest in video indexing and classification techniques. Human activities are very good cues for indexing and classification in most of video sequences. Fig. 1 shows some examples of human movements. As can be seen in these pictures, in most cases, human activities can be described by the motion trajectories generated from body parts suggesting that motion trajectories could potentially be used as features for activity recognition.

This paper presents a new method for extracting motion trajectories from human motions and also shows how to extract temporal segmentation of specific activities from continuous video sequences.

Motion trajectories have several advantages over other features such as intensity [31,32], silhouettes [27–29], and contours [33]. Motion trajectories are very compact as each motion is represented by a pixel location with correspondences between two subsequent frames. Since motion trajectories explicitly specify the movements from the body parts, they are very representative and smooth. Finally, motion trajectories are separable, for they are generated from different body parts separately (e.g., in the first sequence in Fig. 1, motion trajectories are generated from the left hand to the right hand and then to the right foot).

In this work, motion trajectories are used as features for achieving temporal segmentation of *specific* human activities from continuous video sequences. In most of the available video sequences, a large number of video segments with different contents are included in a continuous fashion. Therefore, the temporal segmentation of specific contents is a very critical task in video indexing systems.

To extract motion trajectories without any initialization, dominant motions should be detected and tracked. The dominant motion inturn is extracted from articulated motions, i.e., when the whole arm is moving in a hand gesture, the hand which generates the dominant motion will be detected. Previous motion detection algorithms, such as the ones described in [34–37] and motion segmentation algorithms described in [38,39,42] do not use articulated motions, which are not feasible for our purposes.

In many cases of human activities such as dance and sports typically, the whole body movements are captured in video. In these sequences, body parts (hands and feet) generating the motion trajectories appear as small regions, which makes it very difficult to get enough information of the parts to model their shape or their

* Corresponding author. Tel.: +1 8223 13076350.
*E-mail addresses:* jh643.min@samsung.com (J. Min), r1k@cse.usf.edu (R. Kasturi), camps@whale.ee.psu.edu (O. Camps).
[1] This work was done when Junghye Min was with Department of Computer Science and Engineering in Pennsylvania State University doing her Ph.D.

0262-8856/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2008.03.006

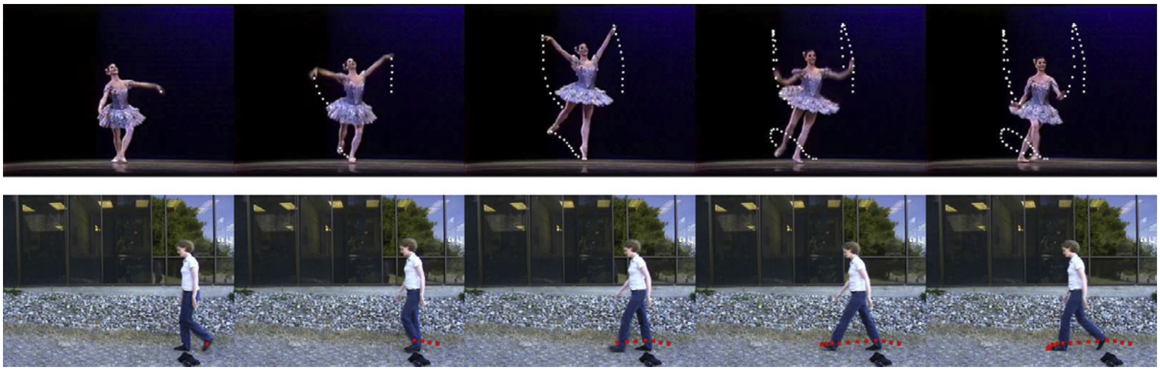

**Fig. 1.** Examples of human activities with motion trajectories extracted using our approach.

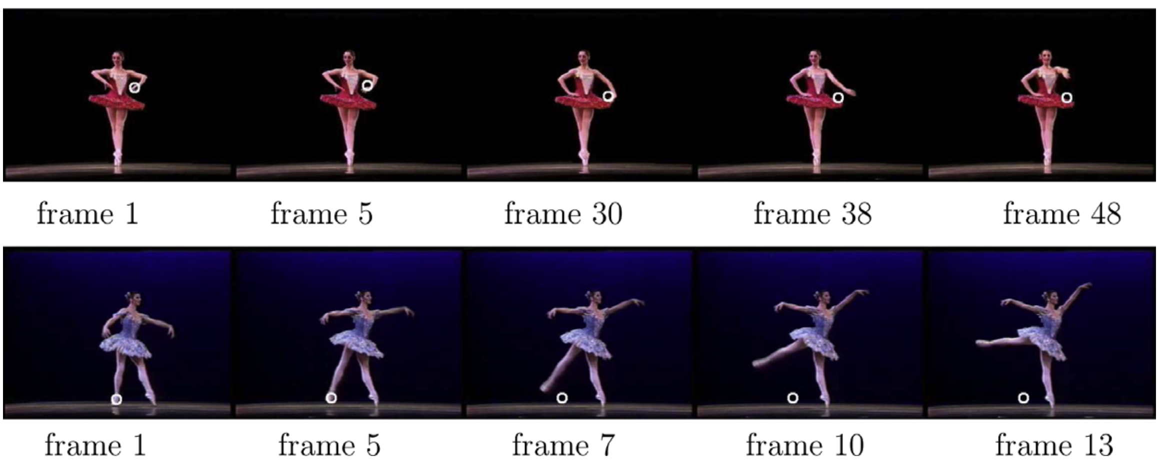


**Fig. 2.** Kernel-based tracking results.

color distribution. Fig. 2 shows the results of the Kernel based object tracking [15] which maximizes the likelihood of the color distribution. For both cases of the hand and the foot tracking, the tracker failed after several frames. Similarity of color distributions between hands and arms (feet and legs) also contributes toward failures.

To overcome the problems described above we propose a new motion segmentation method using mode seeking on the optical flow magnitude to find dominant motion blobs in each frame. The primary significant motion point (SMP) in each motion blob is obtained as a by-product of the motion segmentation algorithm. After the SMPs (Fig. 3) are obtained in every frame, they are used as candidate locations of trajectories, making the tracking procedure possible without any initialization. In each frame, the SMPs are either connected to continuing tracks (trajectories) or new tracks are started from these points. To make the tracking procedure more robust and reliable, our color–optical flow based tracker is applied to each continuing track. This tracker calculates the displacement (tracking result) of continuing tracks in the current frame. This displacement and the SMPs are then used as candidate locations in the current frame, after which the best matches between the continuing tracks and the candidate locations are found by optimizing a cost function.

The multiple motion trajectories obtained by the approach described above are used for temporal segmentation of activities. For each time instance, the optimal alignment between the trajectories from test sequence and the model trajectories is found. A dissimilarity score is then calculated using the dynamic time warp-

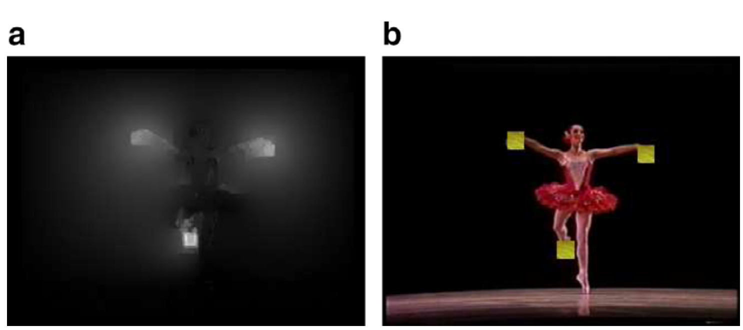


**Fig. 3.** (a) Optical flow magnitude. (b) SMPs.