

Human Activities Segmentation and Location of Key Frames Based on 3D Skeleton

A. Method

0. Ideas

Considering kinematics energy information to recognize the key frame instead of the numerical features. Two kinds of energy is taken into thinking, kinetic energy and potential energy. Intuitively, the frame with a local minimum energy is a key frame in that time interval. Another consideration is about avoiding over-segmentation. From the local minimum energy frame, it iteratively merges the next frame according to a calculated energy similarity and a setting threshold. Therefore, it separates two continuous segmentations. If we see from the energy graph of frame, we can see that a “mountain” is a segmentation.

What's more, the kinetic energy(动能) is calculated using the velocity of joints while the potential energy(势能) is calculated using the position of joints.

// This paper just studies action segmentation instead action recognition.

[Problems to Discuss]

1. How to decide where to do a segmentation in the sequence?
2. How to build an energy model?

B. New things

- 1) Taking energy information to realize action segmentation.
- 2) Pay attention to the over-segmentation problem.
- 3) The accuracy of segmentation is rather high, up to about 95% in average.
- 4) Taking human-object potential energy into consideration, but it needs to locate the position of the nearby objects from the cloud RGB image.

C. Shortcomings

- 1) Consider the human-object interaction may make the segmentation too complex to finish the algorithm in a tolerant time. In other words, it may not be able to work in real-time environment.