MDS-Based Multi-axial Dimensionality Reduction Model For Human Action Recognition

Redha Touati and Max Mignotte
DIRO, Département d'Informatique et de Recherche Opérationnelle, P.O. 6128, Montréal (Québec), H3C 3J7
 touatire@iro.umontreal.ca

Introduction

- We want to develop a new unsupervised human recognition system which is able to recognize more than two classes with a low complexity.
- Recognize automatically low-level actions such as running, walking, hand clapping, etc. from an input video sequence.

Prototype Extraction

- Consists in building a set of prototypes or, more precisely, a set of reliable, compact and discriminative representations of the data cube containing the sequence of binarized silhouettes.
- Generated by an MDS based non linear dimensionality reduction technique from different view-points of the video cube, more precisely view-points 1 and 2.
- The first viewpoint aims at reducing the dimensionality of the image cube along the temporal axis of the video sequence by modeling every silhouette image frames in a low 3D dimensional space.
- The second viewpoint aims at reducing the dimensionality of the image cube through the spatial (line or column) axis of the set of 2D binarized silhouettes.

Preprocessing

- Consists in obtaining the binary video sequence of 2D silhouettes for each human action.
- Subtraction the median background from each image of the sequence and have then used a simple thresholding technique with a classical median filter.
- Centering the gravity center of each binarized body silhouettes inside a rectangular fixed size bounding box.

Proposed Solution

- Background subtraction
- Non linear dimensionality reduction technique
- Classification and fusion

Experimental Results

- Validated by using the leave-one-out procedure and the Weizmann human action data-set

Conclusion

- A simple human action recognition system based on two complementary prototypes.
- The sensitivity of the recognition system to each individual viewpoint is 75.8 and 71.3
- The efficiency of our multi-viewpoint based fusion is 92.3
- Give a better representation of the action in low dimension
- Does not exploit a spatio-temporal characteristics of action

References

- M. Blank and L. Gorelick and E. Shechtman and M. Irani and R. Basri, Actions as space-time shapes, Proceedings of the Tenth IEEE International Conference on Computer Vision, ICCV05, pp. 1396–1402, 2005