
Pattern Recognition

吴庭明

51151500059

2016/06/23

Reference paper

- Human Action Recognition using Depth Maps
 - ◆ Vennila Megavannan
 - ◆ Bhuvnesh Agarwal
 - ◆ R. Venkatesh Babu

Roadmap

- Action Representation
 - ◆ Motion history image
 - ◆ Average Depth Image
 - ◆ Depth Difference Image
- Implementation
 - ◆ Tools and dataset
 - ◆ Feature extraction results
 - ◆ Training
 - ◆ Testing

Motion History Image

$$B(i, j, t) = \begin{cases} 1, & \text{if } D(i, j, t) > 0, \\ 0, & \text{otherwise,} \end{cases}$$

$$B_{diff}(i, j, t) = B(i, j, t) - B(i, j, t - 1)$$

$$I_{mhi}(i, j, t) = \begin{cases} \tau, & \text{if } B_{diff}(i, j, t) = 1, \\ \max(0, I_{mhi}(i, j, t - 1) - \tau) & \text{o.w,} \end{cases}$$

$$I_{mhi}^k = I_{mhi}(i, j, k + N - 1)$$

Average Depth Image

$$I_{avg}^k = \frac{\sum_{t=k}^{k+N-1} D(i, j, t)}{\sum_{t=k}^{k+N-1} B(i, j, t)}$$

Depth Difference Image

$$I_{max}^k(i, j) = \max\{D(i, j, t) : D(i, j, t) \neq 0, \\ \forall t \in [k \dots (k + N - 1)]\}$$

$$I_{min}^k(i, j) = \min\{D(i, j, t) : D(i, j, t) \neq 0, \\ \forall t \in [k \dots (k + N - 1)]\}$$

$$I_{diff}^k = I_{max}^k - I_{min}^k$$

Roadmap

- Action Representation
 - ◆ Motion history image
 - ◆ Average Depth Image
 - ◆ Depth Difference Image
- Implementation
 - ◆ Tools and dataset
 - ◆ Feature extraction results
 - ◆ Training
 - ◆ Testing

Information

- Tools

- ◆ Visual Studio 2010
- ◆ Opencv 2.4.10
- ◆ LibSVM 3.2.1

- Dataset

- ◆ Different kind of motion videos
 - ❑ Bend, bowling, box, jump, kick, squat, strength, swim, trumble, wave
- ◆ Each video contains 19 frames
- ◆ Depth Image

Action Representation

- Cut all videos into frames
- Bounding box features
 - ◆ Compute depth difference image(DDI) for each class
 - ◆ Extract 108 features from different windows each of different size

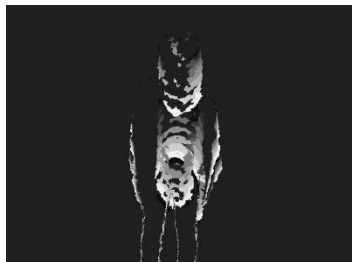
$$F1_k^b = \max_{(i,j) \in R_b} \{I_{diff}^k(i,j) : I_{diff}^k(i,j) \neq 0\},$$

$$F2_k^b = \min_{(i,j) \in R_b} \{I_{diff}^k(i,j) : I_{diff}^k(i,j) \neq 0\},$$

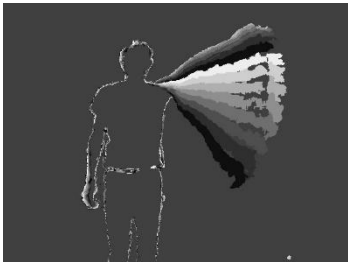
- Hu Moments features
 - ◆ Compute motion history image(MHI) for each class
 - ◆ Compute for each class average depth image(ADI)
 - ◆ Extract 14 features from depth different image

Action Representation

● MHI



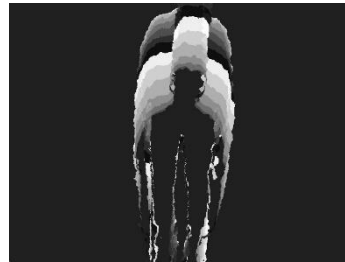
bend



bowling



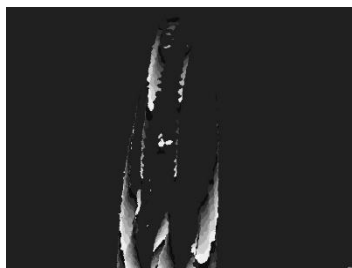
box



jump



kick



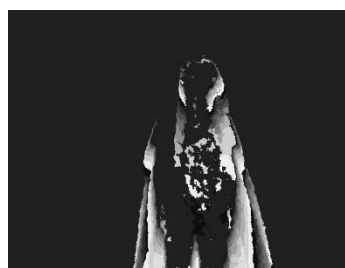
squat



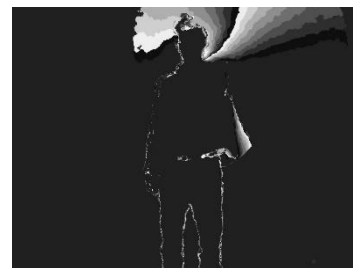
strenth



swim



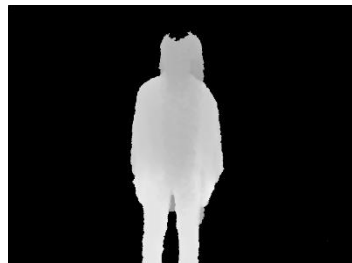
tumble



wave

Action Representation

● ADI



bend



bowling



box



jump



kick



squat



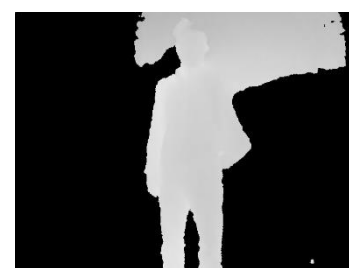
strenth



swim



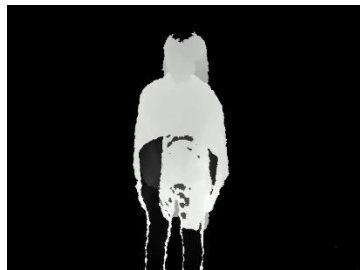
tumble



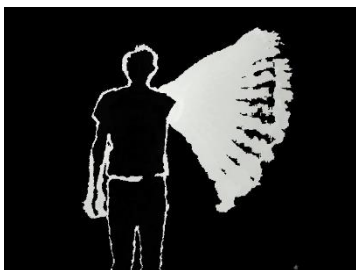
wave

Action Representation

● DDI



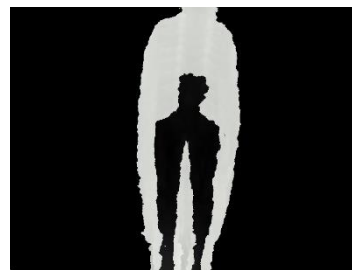
bend



bowling



box



jump



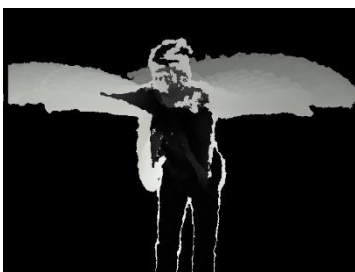
kick



squat



strenth



swim



tumble



wave

Training & Testing

- SVM Type

- ◆ C-SVC (multi-class classification)

- Kernel

- ◆ Linear

- 108(DDI)

- Cross Validation Accuracy = 95.1872%
 - Accuracy = 100% (20/20) (classification)

- 14

- Cross Validation Accuracy = 23.5294%
 - Accuracy = 25% (5/20) (classification)

Training & Testing

- SVM Type

- ◆ C-SVC (multi-class classification)

- Kernel

- ◆ polynomial

- 108(DDI)

- Cross Validation Accuracy = 94.1176%
 - Accuracy = 100% (20/20) (classification)

- 14

- Cross Validation Accuracy = 10.1604%
 - Accuracy = 10% (2/20) (classification)

Training & Testing

- SVM Type

- ◆ C-SVC (multi-class classification)

- Kernel

- ◆ RBF

- 108(DDI)

- Cross Validation Accuracy = 20.3209%
 - Accuracy = 20% (4/20) (classification)

- 14

- Cross Validation Accuracy = 22.9947%
 - Accuracy = 25% (5/20) (classification)

Training & Testing

- SVM Type

- ◆ C-SVC (multi-class classification)

- Kernel

- ◆ 108(MHI)

- ☐ Linear

- Cross Validation Accuracy = 97.3262%
 - Accuracy = 95% (19/20) (classification)

- ☐ Polynomial

- Cross Validation Accuracy = 96.7914%
 - Accuracy = 95% (19/20) (classification)

Training & Testing

- SVM Type

- ◆ C-SVC (multi-class classification)

- Kernel

- ◆ 108(ADI)

- ☐ Linear

- Cross Validation Accuracy = 10.1604%
 - Accuracy = 10% (2/20) (classification)

- ☐ Polynomial

- Cross Validation Accuracy = 10.1604%
 - Accuracy = 10% (2/20) (classification)