Abstract

Stroke is a leading cause of long-term adult disability. Stroke patients can recover through rehabilitation programs prescribed by occupational therapists (OT); however, an individualized rehabilitation program can reduce recovery times compared to traditional ones. We propose a daily activity observation system (DAOS) that uses a Kinect v2 sensor to collect and retrieve motion data. The depth information allows us to perform action recognition and assessment. Action recognition accuracy is 97% on a multi-class kitchen action dataset. Combining an action and its opposite improves recognition rates. Assessment on critical metrics such as arm extension, mean velocity, and max velocity from depth sequences, in 2013 IEEE Conference on Computer Vision and Pattern Recognition, June 2013, pp. 716–723.

Introduction

- 795,000 suffer from a stroke in the U.S. [1]
- The Kinect is a cheap, accurate camera sensor
  - Color camera
  - Depth information
  - Skeleton data
- Daily Activity Observation System (DAOS)
  - Action Recognition
  - Assessment
- No current system integrates action recognition and assessment!

Dataset

- We collected an in-house dataset consisting of common kitchen tasks, as in Table 1 and Fig. 2.
- 9 people x 28 actions x 2 = 504 total.

<table>
<thead>
<tr>
<th>Dataset Name</th>
<th>Actions</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>PickPutDataset1</td>
<td>PickUpCounter</td>
<td>37.5%</td>
</tr>
<tr>
<td>PickPutDataset2</td>
<td>PickUpCounter &amp; PutDown</td>
<td>69.4%</td>
</tr>
<tr>
<td>OpenCloseDataset1</td>
<td>OpenTopCabinet</td>
<td>54.2%</td>
</tr>
<tr>
<td>OpenCloseDataset2</td>
<td>OpenTopCabinet &amp; Close</td>
<td>75.0%</td>
</tr>
<tr>
<td>MixedDataset1</td>
<td>ManipulateFridge</td>
<td>97.2%</td>
</tr>
<tr>
<td>MixedDataset2</td>
<td>WashSink</td>
<td>97.2%</td>
</tr>
</tbody>
</table>

Table 1. Chosen kitchen tasks for our dataset.

Action Recognition

- We use histogram of oriented 4D normals (HON4D) to compute a video descriptor [2].
  - Normals are computed as: \[ \mathbf{n} = \mathbf{V}_{S} = \left( \frac{\partial z}{\partial x} \frac{\partial z}{\partial y} \frac{\partial z}{\partial t} - 1 \right)^T \]
  - The bins of the histogram are set as the vertices of a polychoron, \( F = \{|p|\} \). Bin contributions are calculated as
  - After computation of HON4D descriptors, they are normalized.

Assessment

- Major assessment:
  - Distance:
    - \[ d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \]
  - Speed:
    - \[ r = \frac{d}{T} \]
- Graphs of hand symmetry and chest sway
- To account for occlusion, when any of the upper body joints that are used to perform assessment are inferred or missing, all skeleton data for that instance is set to zero.

Recognition Results

- The in-house dataset in Table 1 was subdivided into smaller datasets as in Table 2.
- SVM with a quadratic kernel to classify the data
- Accuracies are highest on datasets whose actions are different categories. The scores are lower when an action and its opposite are included, such as opening a jar and closing it.

Conclusions

- We present a novel solution for occupational therapists to create more personalized care for stroke rehabilitation patients.
- Recognizes several actions with a high degree of certainty using HON4D as a global descriptor
- Assessment on critical metrics such as arm extension, mean velocity, and max velocity
- Allows a therapist to easily see problem areas or improvements over time to better provide care for a stroke victim

References


Acknowledgements
This material is based upon work supported by the National Science Foundation under Award Number: CNS-1659134.