overview
activity pipeline

Raw Signal → Derived Signals → Action

- Staff Tracking
- Patient Tracking
- Equipment ID
- Gestures
- Binary Events
- i.e. “Greeting”
activity pipeline

Raw Signal → Derived Signals → Action

- Staff Tracking
- Patient Tracking
- Equipment ID

Gestures

Today’s Focus

i.e. “Greeting”

Binary Events
Action Recognition

Staff gestures
- Determine sub-actions
  - i.e. Inserting tube
- Correlate input w/ each known gesture’s PCA basis
- Temporal window?

Activity Logger
- Fuse all derived signals and output “action”
- Time-series graphical model (i.e. CRF, S-LDS, HMM)

Patient Body Tracker
- Establish metric for “unsafe” motion
- Track body over time
- Trigger alert (and log) when problematic
Authors: Ali Bigdelou, Tobias Benz, Loren Schwarz, Nassir Navab

1) Simultaneous Categorical and Spatio-Temporal 3D Gestures Using Kinect, *(3DUI’12)*

2) An Adaptive Solution for Intra-Operative Gesture-based Human-Machine Interaction, *Intelligent User Interfaces (ICIUI’12)*

3) Learning Gestures for Customizable Human-Computer Interaction in the Operating Room, *MICCAI’11*
Learn and identify a set of continuous gestures for manipulating medical imagery

[Bigdelou, MICCAI’11]
sensing modalities

Kinect

IMUs

[Bigdelou, MICCAI’11]
1) Lots of data!

- **Kinect:** 15 joints * (X,Y,Z) = 45 dimensions
- **IMUs:** 4 devices * Quaternion = 16 dimensions

**Solutions:**

- Dimensionality reduction:
  - PCA + Max Likelihood
  - Manifolds + Particle Filter + Max Likelihood

2) Phase detection

- Normalized spatial value

**Solution:**

- Kernel Regression Map

[Bigdelou, 3DUI'12]
Principal Components Analysis

Intuition: Find the greatest variance in the data!
dimensionality reduction

All gesture poses → Per class PCA component → Per class normalization
dimensionality reduction

[Bigdelou, MICCAI’11]
kernel regression

Intuition: weight test data with each of the training data

Calculate normalized value:

\[ \hat{x}_t = f(s_t) = \sum_{i=1}^{n} \frac{w_i(s_t)}{\sum_{j=1}^{n} w_j(s_t)} \cdot x_i. \]

Calculate weights:

\[ w_i(s_t) = k(s_t, s_i) = \exp\left(-\frac{1}{2}\| (s_t - s_i)/\sigma \|^2 \right) \]

Calculate class:

\[ \hat{c}_t = c_{k_t}, \quad \text{where} \quad k_t = \arg\max_i w_i(s_t). \]

Smooth weights

\[ \tilde{w}_i(s_t) = w_i(s_t) \cdot \exp\left(-\frac{1}{2}\left( (i - k_{t-1})/\sigma \right)^2 \right). \]
Manifold learning

Intuition: Find linear embedding from non-linear space
**dimension reduction**

**Laplacian Eigenmaps** (16 dim to 2 dim)

1) Find nearest neighbors
2) Calculate “commute” time between each neighbor
3) Compute eigenvalues of commutes

[Bigdelou, MICCAI’11]
kernel regression

Intuition: weight test data with each of the training data

Calculate normalized value:

\[ \hat{x}_t = f(s_t) = \sum_{i=1}^{n} \frac{w_i(s_t)}{\sum_{j=1}^{n} w_j(s_t)} \cdot x_i. \]

Calculate weights:

\[ w_i(s_t) = k(s_t, s_i) = \exp(-\frac{1}{2}\| (s_t - s_i)/\sigma \|^2) \]

Calculate class:

\[ \hat{c}_t = c_{k_t}, \quad \text{where} \quad k_t = \arg \max_i w_i(s_t). \]

Smooth weights

\[ \tilde{w}_i(s_t) = w_i(s_t) \cdot \exp(-\frac{1}{2}((i - k_{t-1})/\sigma)^2). \]
Particle filter

Intuition: Estimate class from sampling of prior distributions

[Bigdelou, MICCAI’11]
Particle filter

\[ p(s_t | c_t^i, x_t^i) \propto \mathcal{N}(g_{c_t}(x_t^i); s_t, \text{cov}(S_{c_t}^i)) \mathcal{N}(f_{c_t}(s_t); x_t^i, \text{cov}(X_{c_t}^i)) \]

Then:

- Calculate max likelihood of classes
- Average \( x \) over estimated class

<table>
<thead>
<tr>
<th>Prob of sensor reading for each class/embedding</th>
<th>Sample current embedded value centered at each training point (pose space)</th>
<th>Sample current pose value centered at training point (embedded space)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x = \text{embedding} )</td>
<td>( s = \text{pose} )</td>
<td>( g() : x \rightarrow s )</td>
</tr>
<tr>
<td></td>
<td>( f() : s \rightarrow x )</td>
<td></td>
</tr>
</tbody>
</table>
pipeline

[Bigdelou, 3DUI’12]
kinect results

[Bigdelou, 3DUI’12]
imu results

[Bigdelou, MICCAI’12]
user study

[Bigdelou, 3DUI’12]
### user study

<table>
<thead>
<tr>
<th></th>
<th>a)</th>
<th>b)</th>
<th>c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pragmatic quality (PQ)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hedonic quality (HQ)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>too self-oriented</td>
<td>P Mean</td>
<td>P Mean</td>
<td>P Mean</td>
</tr>
<tr>
<td>self-oriented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>neutral</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>task-oriented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>too task-oriented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>superfluous</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>desired</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>desired</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Kinect**
  - Common Data Set
  - Personalized Data Set
- **IMU**

[Bigdelou, 3DUl’12]
Questions?