Customizable Gesturing Interface
for the Operating Room using Kinect

CVPR 2012 Gesture Recognition Workshop, Providence, Rohde Island, USA

Ali Bigdelou, Tobias Benz, Loren Schwarz, Nassir Navab
Computer Aided Medical Procedures (CAMP)
Technical University of Munich, Germany

Technische Universität München

Computer Aided Medical Procedures

Motivation

Interaction with computerized systems in the OR is challenging for surgeons due to sterility requirements and the complexity of medical procedures

Often a less-skilled assistant controls computers using keyboard and mouse, guided verbally by the surgeon

This indirection can be inefficient, cause misunderstandings and surgeons do not have immediate control

Contribution

We propose a method that allows surgeons to interact with medical systems by means of fully customizable gestures. Microsoft Kinect allows capturing skeletal models. Gestures are learned from a single training example

Exploiting proper feature extraction methods the same training data can be used for different users. The gestures are simultaneously categorical and spatio-temporal [1]

Gesture Recognition Method [2]

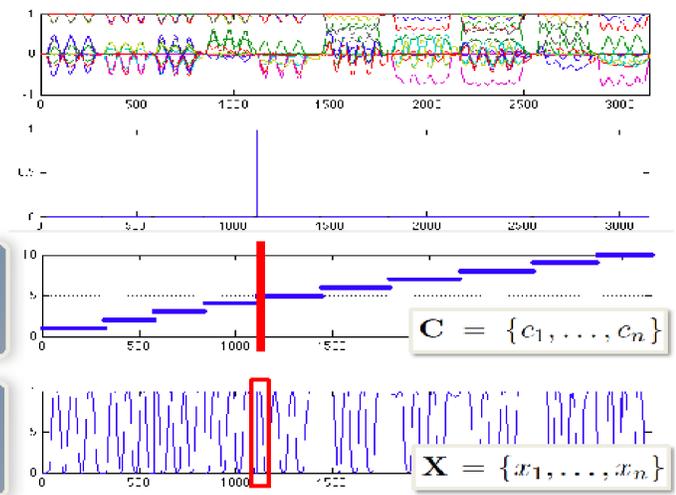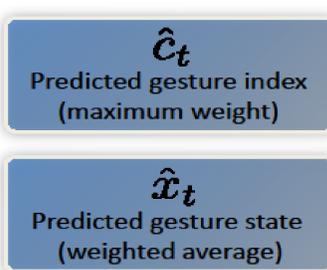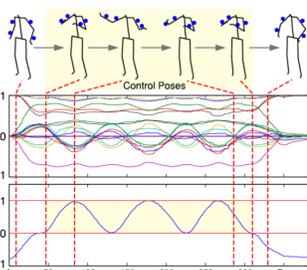Record all the gestures and perform a dimensionality reduction technique such as PCA or Manifold to reduce the recorded data to 1D space and map it to [0,1]

Concatenate the recorded sample data and create two additional sets one for dimensionality reduced data and one for assigned indices for each trained gesture

Using the kernel regression compute the index of the current gesture and predict the user state within that gesture

$$\hat{x}_t = f(\mathbf{s}_t) = \sum_{i=1}^{n} \frac{w_i(\mathbf{s}_t)}{\sum_{j=1}^{n} w_j(\mathbf{s}_t)} \cdot x_i$$

$$w_i(\mathbf{s}_t) = \exp(-\tfrac{1}{2}\|(\mathbf{s}_t - \mathbf{s}_i)/\sigma\|^2)$$

$\hat{c}_t$
Predicted gesture index (maximum weight)

$\hat{x}_t$
Predicted gesture state (weighted average)

$\mathbf{C} = \{c_1, \ldots, c_n\}$

$\mathbf{X} = \{x_1, \ldots, x_n\}$

Flexible and Customizable Implementation [3]

Our system enables an easy and dynamic association of learned gestures to the properties of arbitrary systems

To freely customize the behavior of the system for a specific scenario, we implemented the Signal.NET framework based on component model and data pipeline design pattern

A visual editor has been developed in order to facilitate the customization process without any programming
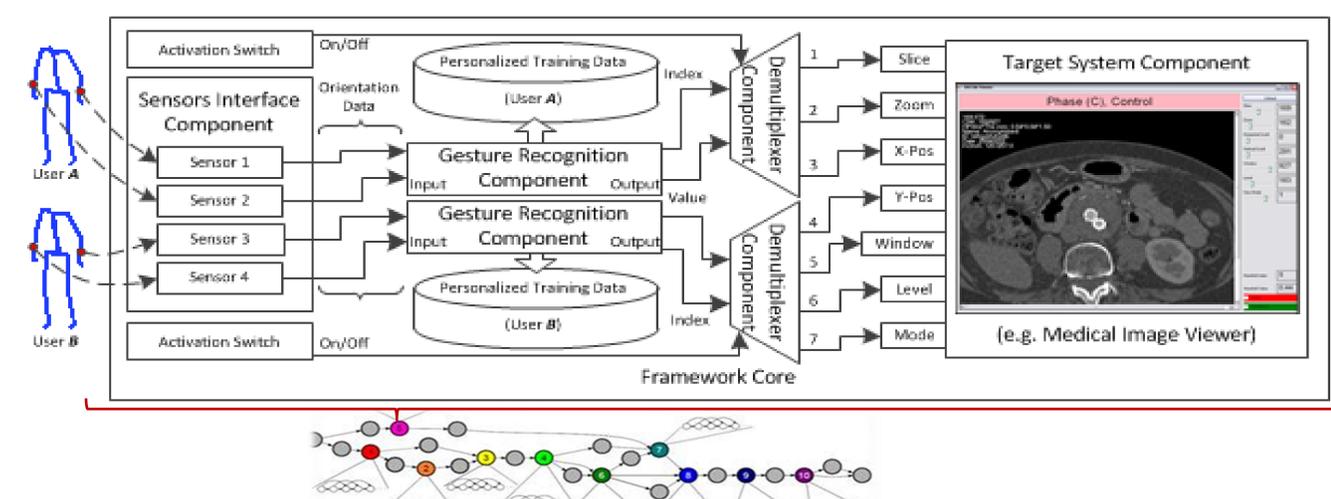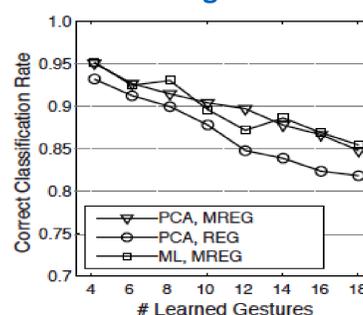
Experiments and Results

Experimental Setup:

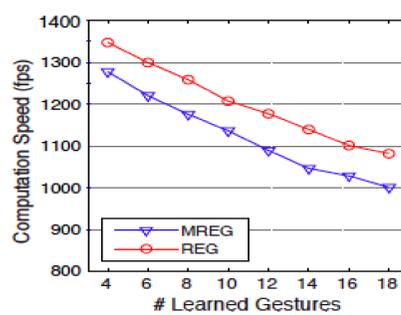User study with more than 30 test subjects, including surgeons, UI experts

Test tasks were designed based on real intra-operative scenarios

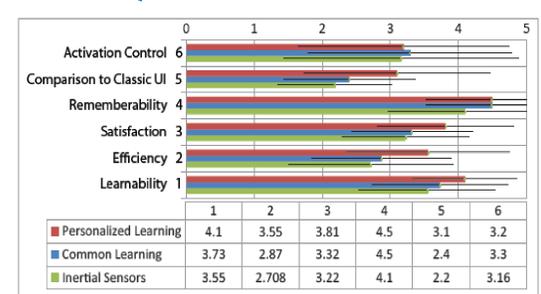Wide range of quantitative and qualitative properties are explored and reported [1, 2, 3]

Correct Recognition Rate

Performance

Qualitative User Studies

	1	2	3	4	5	6
Personalized Learning	4.1	3.55	3.81	4.5	3.1	3.2
Common Learning	3.73	2.87	3.32	4.5	2.4	3.3
Inertial Sensors	3.55	2.708	3.22	4.1	2.2	3.16

Activation Control 6
Comparison to Classic UI 5
Rememberability 4
Satisfaction 3
Efficiency 2
Learnability 1

References

1. A. Bigdelou, T. Benz, L. Schwarz and N. Navab. **Simultaneous Categorical and Spatio-Temporal 3D Gestures Using Kinect**. *International IEEE Symposium on 3D User Interfaces (3DUI)*, Orange County, CA, USA, March 2012.
2. A. Bigdelou, L. Schwarz and N. Navab. **An Adaptive Solution for Intra-operative Gesture-based Human-Machine Interaction**. *International ACM Conference on Intelligent User Interfaces (IUI)*. Lisbon, Portugal, February 2012.
3. A. Bigdelou, L. Schwarz, T. Benz and N. Navab. **A Flexible Platform for Developing Context-Aware 3D Gesture-based Interfaces**. *International ACM Conference on Intelligent User Interfaces (IUI)*, Lisbon, Portugal, February 2012.