Professional Natural Interfaces for Medicine Applications

Illya Starodubtsev[1], Vladimir Averbukh[1],
Nataly Averbukh[2], and Dmitriy Tobolin[3]

[1] Krasovskii Institute of Mathematics and Mechanics, Ural Branch Russian Academy of Sciences, Yekaterinburg, Russia
[2] Urals Federal University, Yekaterinburg, Russia
[3] AngioSystems Ltd., Russia

Abstract. This article focuses on the problems of development of control systems and medical equipment medical professional interfaces. In our paper we propose a solution to the problem, using gesture language and contact-less motion capture.

Keywords: NUI, natural interfaces, professional interfaces, gesture, depth map, medical equipment, control.

1 Introduction

Our interest lies in the field of development of interfaces for professional work. Including work with the medical equipment. It means control of the medical robot, control during diagnostics and the operation, remote control (teleoperation).

1.1 Problem

During deal with the equipment it is necessary to consider the following things.

1. **Sterility**. During the surgery, the doctor is obliged to keep sterility mode. It leads to restriction of use of hardware tools (button, stick, arm, etc.). As well as a passive / active motion capture markers on the body of a surgeon.
2. **Specialized activity**. These interfaces are not oriented to mass users, so it should use a specialized activity.
3. **Quality assessment**. There is a problem of an assessment of quality of such natural interfaces. In our opinion the usual quality assessment methods are not applicable here. In the case of professional interfaces should be considered, in addition estimates of the rate and accuracy of action, a stress level during user activity. To assess stress is possible through surveys and interviews, and by objective measurements.

C. Stephanidis (Ed.): HCII 2014 Posters, Part II, CCIS 435, pp. 435–439, 2014.

1.2 Overview

Research in this field is making steady headway. There are very interesting published over the past two decades. The close problem statements are contained in [1,2,3,4,5]. In these papers a way to assess the quality of the medicine sterile interfaces is given. However, a number of parameters are controversial (for example measuring objectivity and representativeness). On our opinion these approaches are needed to be determined more precisely.

Most of works has a number of drawbacks

1. **Pose vs trajectory**. In the great majority of papers the final control gesture is based on the (carpal) pose without hand trajectory in space. Or based only on the trajectory, without regard (carpal) pose.
2. **Binding to the hardware**. Algorithms and methods is rigidly bound to a specific technology.
3. **Non-flexibility**. In most cases, there are no opportunities to customize and make changes when conditions change or tuning to a specific operator/problem.

1.3 Problem

It is necessary to develop a system of gesture control for medical equipment. It is necessary to

1. Consider observations above described.
2. To design a flexible and convenient system.
3. To design a reliable system.
4. Consider the specifics of the doctor-operators work.

To consider the work specifics were analyzed real activity of the doctor during surgery and diagnostics [6]. Based on the analysis conclusions are drawn:

1. Control gestures should contain information about the pose and spatial information. Information about a pose a lesser extent, about the spatial trajectory more.
2. Gesture language should contain the native fuses which protect against "random" movements. They are of different nature for different types of manipulation.
3. Using virtual instruments such as real analogues, reduces the barrier to entry and increase the stability of the operator work.
4. To control the obsolete equipment designed for "classical" input devices (mouse and keyboard) is working well the idea of *touch-less-screen* + simple gesture language based on analysis of the *point-of-touch*.

2 Techniques Used

In the operating room conditions to use active markers for motion capture or passive markers on the body of doctor is not allowed. Also, the use of color information leads to complication of algorithms (compensation lighting conditions, suppression of active shadow). This leads to a drop in reliability.

To resolve this we will use information about the depth of the scene.

Definition 1. *Will be called **depth map scene** point cloud. Such that each point contains the distance from the origin point of the ray intersects with an object in the scene of a spherical coordinate system or infinity if there is no possibility to calculate the distance.*

$$DepthMap_{Spherical} = \{(\theta, \phi, \rho) : \theta, \phi \in [0, 2\pi], \rho \in \mathbb{R} \cup \infty, \rho = dist(O, P)\}, \quad (1)$$

where O is the origin point and P is an object point.

Obviously, that can be easily represented as a cloud of points in the Cartesian coordinate system:

$$DepthMap_{Cartesian} = \{(x, y, z) : x, y, z \in \mathbb{R}^3 \cup \infty\}. \quad (2)$$

2.1 Pose and Trajectory

Using both methods of processing (carpal) pose and trajectory (position) hand in space produces a favorable result. To use this the system modifiers was developed.

Definition 2. *We will call the **modifier** function having at the input state of the system and enabling the output to logical true if the state of the system satisfies certain conditions, and false otherwise.*

$$Mod : S \longmapsto \{TRUE, FALSE\}, \quad (3)$$

where S is system state.

Notes and Comments. There may be different types of modifiers. From simple button type or manipulation area (the trajectory of interest is treated as a point in the area and is ignored otherwise). To the complex, such as having recognized carpal complex poses or initiating/final gestures. Modifiers affect the interpretation of the results of the analysis of the trajectory.

Using a system of modifiers can be easily (naturally) used path analysis and carpal pose together. And not only.

Through this, it can be easily to build complex gesture languages suitable, comfortable and applicable for a specific task or for a specific person.

2.2 Full-Body and Point Based Interaction

There are two basic approaches to capture the movement to create a gesture language. It is:

1. **Full-Body Based Interaction.** In this case there is recognition full poses a human operator. This provides advantages like:

- It is allow to use information about the position of the set of points, such as hands, feet, elbows, torso, head, etc.
  For example, it is easy to make a pointing gesture as a ray from the elbow in the direction of the carpal.
- With contextual information about the position of other parts of the body in space, easily identifiable point of interest, if more than one. Even with partial (or full) overlap.
  For example, it is always possible to distinguish the right from the left hand.
- Using the information on the form (length, cross-sectional radius) body parts is not difficult to recover the trajectory of point of interest in the partially overlapping the visible part of the limb.

However, this method requires more complex algorithms and more computing resources. Also, it is not efficient when it want to recognize a non-regular (non-standard) or non-humanoid object.

2. **Point Based Interaction.** In this case there is capturing and tracking of the point of interest without reference to the context. This provides advantages like:
   - It is allow to tracking an irregular skeleton. For example, the operator has a big instrument in the hands.
   - It is allow to tracking an arbitrary object. For example, it may be the end of instrument or pointers.
   - Low computational cost.

   However, to identify points need to use the additional information. For example, the proximity of the position given the history. This can be problematic when point is overlapped, or the distance between two points is low.

As can be seen, both approaches have their advantages and are useful in certain situations. To avoid selecting implemented both approaches.

"Full-Body" Approach. To track a cloud of points in the space corresponding to the operator projected on the plane of observation. Projection is multi-surface flat figure with self-intersections. For the projection procedure is applied skeletization. Next, create a master skeleton. And for the obtained (test) and master skeleton solved the problem of minimizing the energy of the residual by using a hierarchical representation. More in [7]. This approach allows to calibrate by one pose and be adjusted to suit various user settings.

"Point" Approach. To track is used analysis of the amount of motion in the frame. This is done using the cumulative difference buffers. More in [8]. This approach has complexity $O(n)$ on the number of operations and required memory and allows to track an arbitrary object. Point of interest initiated by special gesture.

2.3 Binding to the Hardware

In the case of tight binding to a specific hardware reduces overall system reliability. This makes difficult the timely updating of equipment or quick replacement in case of its failure. Especially when the equipment becomes obsolete.

In the case of the possibility to use different algorithms is increasing a flexibility and applicability. This allows to configure to specific user and a specific task. Which increases the efficiency and reliability.

To provide flexibility modular architecture introduced. Hardware module carried out of the kernel. This allows you to maintain different types of depth sensors by writing a plugin - reader/converter (like a driver). Now, there are plugins for some types of lidar, RGBD cameras, stereo cameras, TOF and rangefinders.

As well as algorithms. Analysis module can be replaced with a more appropriate without changing the kernel. In combination with the modifier system it is give excellent result for flexibility to the system.

3 Prospects

As near term prototypes of interfaces for real medical devices are considered. In particular prototype design gesture interface for control of Angiographic Systems conducted.

Another direction of our research and development is related to the use of gesture and motion capture interface to monitor and create an accurate log for modern cardio surgeries.

References

1. Wachs, J.P., Stern, H.I., Edan, Y.: Parameter search for an image processing fuzzy C-means hand gesture recognition system. In: Proc. of IEEE Intl Conference on Image Processing, ICIP, Spain, vol. 3, pp. 341–344 (2003)
2. Stern, H.I., Wachs, J.P., Edan, Y.: Human Factors for Design of Hand Gesture Human-Machine Interaction. In: IEEE International Conference on Systems, Man and Cybernetics, SMC 2006, vol. 5, pp. 4052–4056 (2010)
3. Wachs, J.P.: Gaze: Posture and Gesture Recognition to Minimize Focus Shifts for Intelligent Operating Rooms in a Collaborative Support System. International Journal of Computers Communications & Control 5(1), 106–124 (2010)
4. Kristensson, P.O., Nicholson, T.F.W., Quigley, A.: Continuous Recognition of One-Handed and Two-Handed Gestures using 3D Full-Body Motion Tracking Sensors. In: Proceedings of the ACM International Conference on Intelligent User Interfaces, pp. 89–92 (2012)
5. Jacob, M., Wachs, J.P.: Context-based Hand Gesture Recognition for the Operating Room Pattern Recognition Letters (2013) (article in press)
6. Averbukh, V., Starodubtsev, I., Tobolin, D.: The Gesture Interface for Control of Angiographic Systems. Modern computer technology. In: Proceeding of International Russian-Korean Conference, pp. 97–107. UrFU, Ekaterinburg (2012)
7. Starodubtsev, I.: Motion captire based on flexible skeleton hierarchy. Scientific and Technical Bulletin of the Volga Region 1, 159–162 (2014)
8. Starodubtsev, I.: Depth map based initialized motion capture and tracing for gesture interface. Scientific and Technical Bulletin of the Volga Region 3, 264–268 (2013)