A Wearable Visuo-Inertial Interface for Humanoid Robot Control

Junichi Sugiyama
Dept. of Computer Science and Engineering
Toyohashi University of Technology
Toyohashi, Aichi, Japan
Email: sugiyama@aisl.cs.tut.ac.jp

Jun Miura
Dept. of Computer Science and Engineering
Toyohashi University of Technology
Toyohashi, Aichi, Japan

***Abstract*—This paper describes a wearable visuo-inertial interface for humanoid robot control, which allows a user to control motion of a humanoid robot intuitively. The interface composed of a camera and inertial sensors and estimates body motion of the user: movement (walk), hand motion, and grasping gesture. The body motion (walk) estimation is performed by a combination of a monocular SLAM and a vision-inertial fusion using an extended Kalman filter. The hand motion is also estimated by using the same motion model and sensor fusion as the body motion estimation. The estimated motion was used to operate the movement or arm motion of the humanoid robot. We conducted the experiment on robot operation. The results revealed that the user intuitively controlled the robot and it responded to the operator commands correctly.**

***Keywords*—human robot interaction; humanoid robots**

I. Introduction

Service robots are expected to support human in various everyday situations. However it is difficult to give a robot a complete set of required skills and knowledge to achieve tasks in advance. On-site teaching is an effective way to transfer the skills and the knowledge to the robot.

In this paper, we adopt human body motion to teach robot motion as a natural and intuitive method. Such a method for supervising the robot's motion can be realized through capturing the human motion. There are several methods for measurement of human motion: by motion capture system, by environmental cameras [1], or through attachment of many sensors to the user's body [2] [3]. These approaches have several limitations such as narrow field of view and equipment that is difficult to wear.

For estimation of user's body motion we developed a wearable interface with a camera and two inertial sensors. The information from these sensors is used to estimate body movement (walk) and hand motion of the user. The estimated body movement and hand motion is used to operate the movement and arm motion of a humanoid robot. We implemented an experimental system and successfully applied it to controlling the humanoid robot.

II. System Overview

The robotic system is composed of the body motion estimation interface and the mobile humanoid robot system.

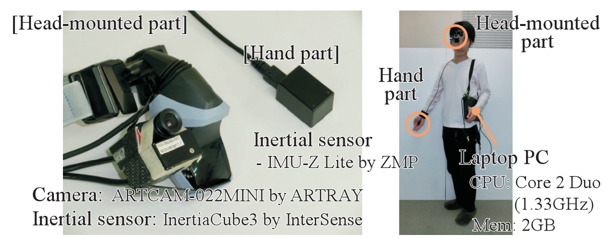
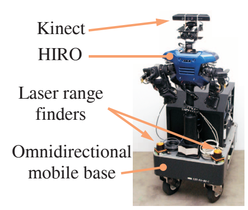
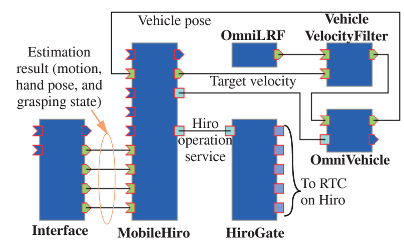


Fig. 1. User with the body motion estimation interface.



(a) Humanoid robot, a Hiro on an omni-directional mobile base.



(b) System diagram of the robot system using OpenRTM, with the major inputs, outputs, and services.

Fig. 2. Robot and software system.

Fig. 1 shows a user with the proposed interface. The wearable interface consists of the head-mounted part and the hand-worn part. The head-mounted part has a camera and a 9-axis orientation sensor attached to a monocular head-mounted display. The orientation sensor includes a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer. The hand-worn part has an inertial sensor with a 3-axis accelerometer and a 3-axis gyroscope.

Fig. 2a shows the humanoid robot, Hiro by Kawada Industries on an omni-directional mobile base by KER. In this robotic system, the functional software modules (Fig. 2b) are constructed as RT components (RTC) with OpenRTM (http://www.openrtm.org/). *InterfaceGate* receives the estimated poses of the user's body, the hand motion, and grasping gesture from the interface. Then, this RTC transforms these data into the head tilt angle, the pose of right hand of Hiro, and the pose of the mobile base. These transformed data is sent to *MobileHiro* as target inputs. *MobileHiro* controls Hiro and the mobile base: Hiro is controlled through HiroGate which serves as a gateway to its motion services; the velocity of the mobile base is controlled by *OmniVehicle*. *OmniLRF* and *VehicleVelocityFilter* are used for detection of collision with obstacles by using the laser range finders (Fig. 2a).

III. Body Motion Estimation Interface

The interface (Fig. 1) is based on the body motion estimation method which we have developed [4]. This method combines a monocular SLAM [5] with a vision-inertial fusion using an extended Kalman filter [6] for both movement and hand motion estimation. By the vision-inertial sensor fusion, a drift-less and robust movement estimation is realized without any environmental sensors.

The previous system used a Wii remote with a visual marker for hand motion estimation. In this system, we estimate the hand motion without any markers. The new method detects the hand in the camera image based on skin color and pairs of edges which lie in a predetermined finger width [7]. It can also detect the grasping gesture by analyzing the edge intensity in the hand region (Fig. 3c, d).

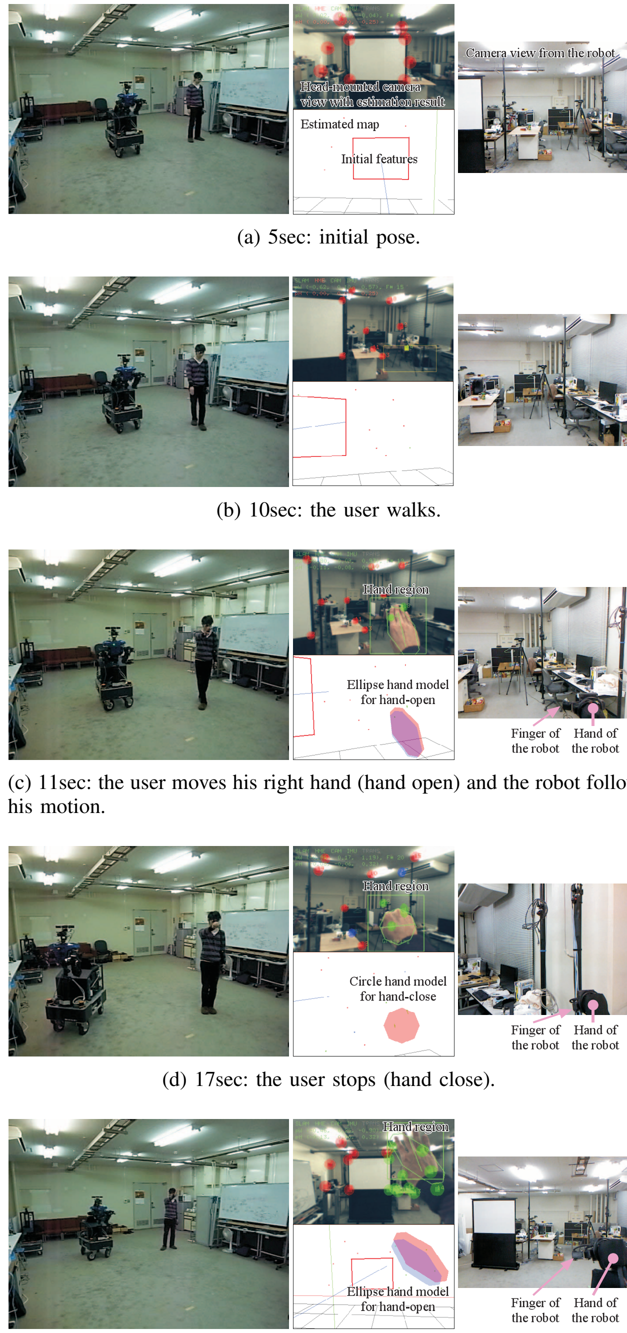
IV. Experimental Result

We carried out an experiment of the robot operation using the interface. Fig. 3 shows the experimental procedure. In this experiment, the user wearing the interface walks and moves his hand simultaneously. The interface estimated the user's motion successfully (see the middle column of Fig. 3) and sent the estimated poses and the grasping gesture to the robot. The robot then followed the user's motion correctly (the left and the right column of Fig. 3).

V. Conclusions

We proposed a wearable visuo-inertial interface for controlling the arm/hand/mobile platform motion of the robot by the human body movement. The interface has a camera and two inertial sensors; the streams of information from those sensors are fused on-line by an EKF for fast and robust body motion estimation. The interface enables us to control a humanoid robot naturally and intuitively without any environmental sensors.

The future work will be devoted to the application of the system to mobile manipulation tasks in, for example, remote-control and/or telepresence applications. This will require various sensory feedback from the remote robot such as visual and tactile information for a safe and easy robot control. Integration of autonomous and manual control modes for performing complex tasks is also the scope of research interest.



(a) 5sec: initial pose.

(b) 10sec: the user walks.

(c) 11sec: the user moves his right hand (hand open) and the robot follows his motion.

(d) 17sec: the user stops (hand close).

(e) 45sec: the user moves backward and stops (hand open).

Fig. 3. Result of the robot operation experiment.

References

[1] T. Matsuyama, X. Wu, T. Takai, and T. Wada, "Real-time dynamic 3d object shape reconstruction and high-fidelity texture mapping for 3d video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. CSVT-14, no. 3, pp. 393–434, 2004.

[2] T. Harada, T. Gyota, Y. Kuniyoshi, and T. Sato, "Development of wireless networked tiny orientation device for wearable motion capture and measurement of walking around, walking up and down, and jumping tasks," in *Proc. of IEEE/RSJ 2007 Int. Conf. on Intelligent Robots and Systems*, pp. 4135–4140, 2007.

[3] T. Maeda, H. Ando, M. Sugimoto, J. Watanabe, and T. Miki, "Wearable robotics as a behavioral interface – the study of the parasitic humanoid –," in *Proc. of 6th Int. Symp. on Wearable Computers*, pp. 145–151, 2002.

[4] J. Sugiyama and J. Miura, "A wearable robot control interface based on measurement of human body motion using a camera and inertial sensors," in *Proc. of the 2011 IEEE Int. Conf. on Robotics and Biomimetics*, pp. 565–570, 2011.

[5] J. Montiel, J. Civera, and A. Davison, "Unified inverse depth parametrization for monocular slam," in *Proc. of Robotics: Science and Systems*, 2006.

[6] L. Armesto, J. Tornero, and M. Vincze, "Fast ego-motion estimation with multi-rate fusion of inertial and vision," *The Int. Journal of Robotics Research*, vol. 26, no. 6, pp. 577–589, 2007.

[7] N. Petersen and D. Stricker, "Fast hand detection using posture invariant constraints," in *Proc. of the 32nd annual German conference on Advances in artificial intelligence*, pp. 106–113, 2009.