Multi-sensor Real-time Motion Tracking System

YANG-KEUN AHN, KWANG-SOON CHOI, YOUNG-CHOONG PARK
Korea Electronics Technology Institute
121-835, 8th Floor, #1599, Sangam-Dong, Mapo-Gu, Seoul
REPUBLIC OF KOREA
ykahn@keti.re.kr

Abstract: - This study proposed a multi-sensor real-time motion tracking technology based on a Kinect camera and Attitude and Heading Reference System (AHRS) sensors. The proposed method serves as a test bed for sportic human motion analysis. Joint data was retrieved from the Kinect camera, and accurate motion tracking was performed through AHRS sensors attached to major joints. OpenGL contents were generated to support the motion tracking technology, and the output showed that the proposed system is more accurate that other systems that rely on single sensors. The proposed system uses minimal information required for human motion tracking, making it more affordable and easier to set up. The experimental results verified that the system is capable of realistic motion capture.

Key-Words: - Gesture Tracking, Joint Detection, Gesture Recognition, Screen Control System, OpenGL

1 Introduction

Recently, motion capture has become widespread in various fields. It is used not only as content in games, films and medicine, but also in other fields such as education and the military. While the number of motion capture applications has seen a rapid increase, related equipment and systems have not developed as quickly. The types of motion capture, depending on the media, can be classified into mechanical, optical, magnetic, and ultrasonic. The mechanical method, which usually attaches the motion capture device externally, is accurate and responsive [1]. However, the device can be inconvenient to use, and imposes constraints on preparation and measurement. As such, optical and magnetic approaches account for 70 to 80 percent of motion capture today. The optical method [2] is highly accurate but difficult to install. A key challenge is to enhance accuracy through various pre-treatment processes and to use calibration to remove distortion. In other words, a high level of accuracy can be attained if problems such as image distortion, input speed, and interference are resolved. The optical method obtains measurements in all directions, improving accuracy, but real-time processing becomes difficult due to the vast amount of data. Since this method involves extensive system resources, it is difficult to apply to various fields. The disadvantages of the magnetic method [3] include the need for device calibration and relatively poor accuracy. A hybrid method [4] has been proposed to maximize the strengths of the optical method while compensating for its weaknesses. The hybrid method, which aims to develop a low-cost compact motion capture system, uses a Position Sensitive Detector (PSD) and measures position in three-dimensional space with markers instead of cameras. This is an improvement over the conventional method, but users are still burdened by the preparation process and system size.

Today, research on human motion capture varies greatly with the choice of sensor, efforts to minimize uncertainty in sensor data, and proper modeling of the human body [5]. While extensive research has been performed using different sensors to detect various types of human motion and to provide a wide range of related services, users have complained about the inconvenience of attaching sensors, and the accelerometer alone is insufficient in capturing diverse movements. In other words, studies that involve single sensors are limited for capturing sportic motion. To resolve this issue, one study combined data from both image sensors and accelerometers for more accurate human motion capture [6].

The multi-sensor real-time motion tracking technology proposed in this study captures human motion by integrating position information from a Kinect camera and direction information from motion sensors; it provides real-time content as output.

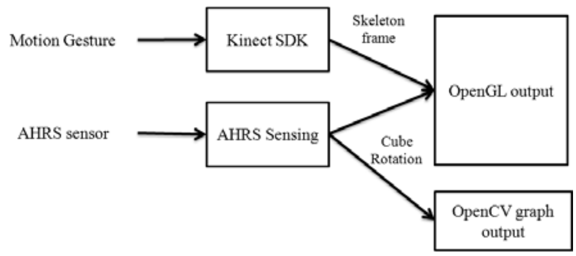

Fig. 1 System flow diagram

2 Kinect Sensor

As shown in Fig. 2, the Kinect sensor consists of three lenses. The lenses on the left and right ends are three-dimensional depth sensors. When the projector shoots infrared as points having pixel units, the depth sensors detect depth based on infrared reflected by the subject via the infrared camera on the right. The camera lens in the middle is a general RGB camera that captures RGB images [7].

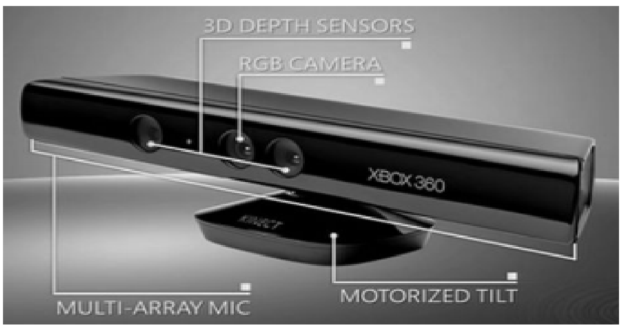

Fig. 2 Windows-based Kinect camera

The range of the Kinect camera is a horizontal angle of 57.5° and a vertical angle of 43.5°. It has a horizontal field of view of 1.2 to 3.5 m, a width of 2.5 m, and a vertical field of view of 1.8 m, which translates to a practical range of 800 to 4000 mm. The device supports VGA images (640 x 480) at 30 fps, and has a skeleton tracking range of 1.2 to 3.5 m. For joint tracking, Kinect recognizes 20 skeletal joints and extracts the three-dimensional information of individual joints.

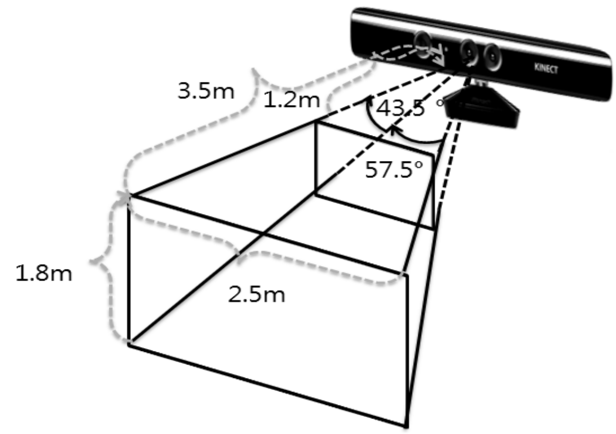

Fig. 3 Range of Kinect camera

3 AHRS Sensor

3DM-GX3-25, which serves as the AHRS system, contains sensors to measure geomagnetism, angular velocity, and acceleration. These sensors are used to calculate the azimuth of the modules [8].

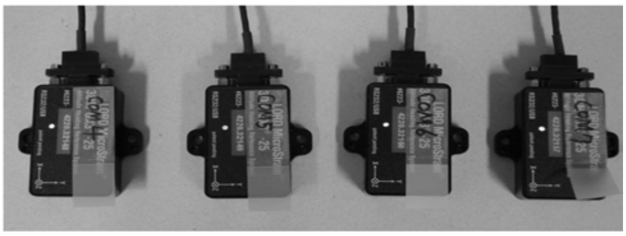

Fig. 4 AHRS sensors

For sensor control, commands in the form of hexadecimal numbers are transmitted to the sensors, and such commands are received as parameter values through the USB/serial port. This is used to retrieve various parameters such as acceleration, geomagnetism, angular velocity and internal module information. The AHRS system of this study extracted Euler angles using the command structure shown in Fig. 5.

Function:	The 3DM-GX3®-25 will output a data record containing Euler Angles.	
Command:		
Byte 1	0xCE	
Response:		
Byte 1	0xCE	
Bytes 2-5	*Roll*	*(IEEE-754 Floating Point)*
Bytes 6-9	*Pitch*	*(IEEE-754 Floating Point)*
Bytes 10-13	*Yaw*	*(IEEE-754 Floating Point)*
Bytes 14-17	*Timer*	
Bytes 18-19	*Checksum*	

Fig. 5 Command structure for Euler angles of sensors

4 OpenGL

OpenGL, developed by Silicon Graphics, is a standard API for two-dimensional and three-dimensional graphics. Comprised of more than 250 function calls, the interface is capable of generating simple geometrical shapes and complex three-dimensional scenes [9].

Joint data extracted from the Kinect camera must be calibrated and normalized for the coordinate system in OpenGL. The x-, y-, and z-axis values of the joint data were calibrated between –1 to 1, and the OpenGL camera was appropriately positioned using the model view before the output of contents. Cube-shaped contents were employed to express the angular changes of the AHRS system.

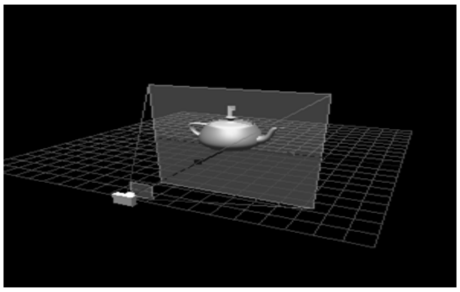

Fig. 6 OpenGL Model View

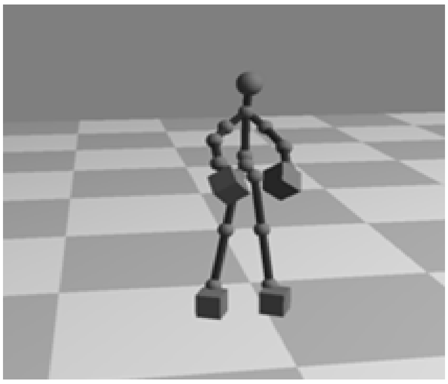

Fig. 7 OpenGL contents

5 System Configuration

To assess the performance of the motion capture content simulation using the proposed multi-sensor real-time motion tracking technology, this study implemented a Hybrid Motion Capture (HMC) system comprised of the Kinect sensor and AHRS sensors. As can be seen from the system configuration in Fig. 8, user input modules detect the movement of the AHRS sensors, and OpenGL contents are the resulting output.

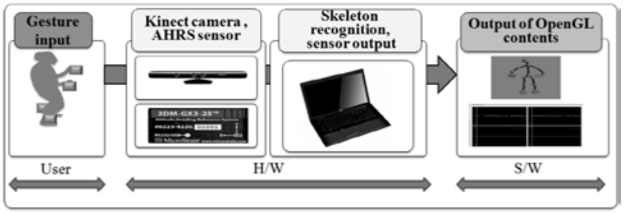

Fig. 8 System configuration

5.1 Hardware Configuration

Fig. 9 shows the hardware setup for HMC; Fig. 10 gives the overall flow block diagram for the hardware. The hardware is this study consisted of a Kinect camera for motion capture, AHRS sensors (Microstrain, 3DM-GX3-25), a PC for data processing, and a display for content output.

Fig. 2 Hardware setup

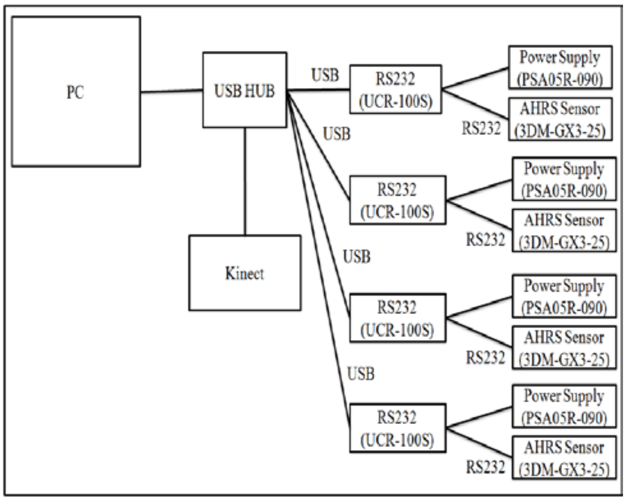

Fig. 10 H/W flow block diagram

5.2 Software Configuration

The GUI of the HMC program consisted of an OpenGL screen for the output of joint data, a Windows screen the output of the RGB and joint information, and a Windows screen for graphs showing AHRS sensors angles. Fig. 11 shows the HMC software flow chart, comprised of a Kinect camera, OpenCV graph output based on AHRS sensor input, and OpenGL content output.

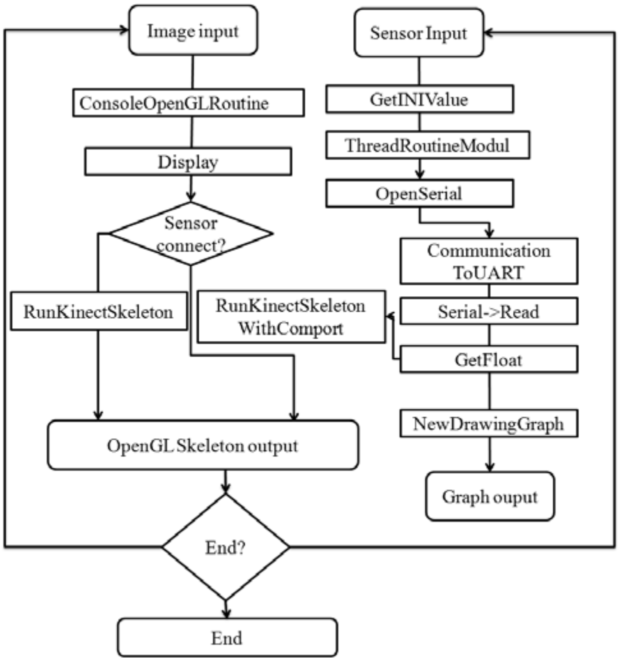

Fig. 3 HMC software flow chart

6 Experimental Results

The software environment was Visual Studio 2010 in Windows 7; the hardware consisted of the AHRS sensor module (3DM-GX3-25), USB/Serial (UCR-100S), switching power supply (PSA05R-090), a computer for processing of input images, and a display for content output. The specifications of the desktop computer were Intel i7-2600K CPU 2.8GHz, 3.48GB. The range of the camera was 1.2 to 3.5 m, and the optimal distance for motion capture was 1.5±0.1m. The resolution was 640x480, and the frame rate for OpenGL contents and motion capture was 30 fps.

The actual measurements obtained using HMC and the content values retrieved as output from the OpenGL were compared to verify the effectiveness of the proposed system.

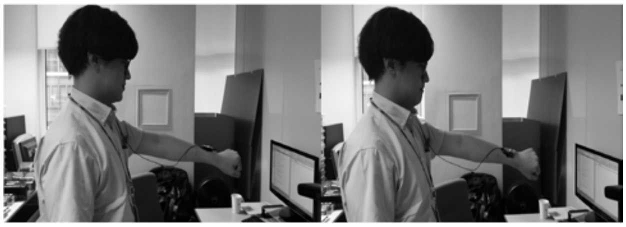

Fig. 12 User participating in the experiment

Table 1 Comparison of measured angle and content angle

Content angle Measured angle	x-axis	y-axis	z-axis
-180°	-178.7°	-178.7°	-178.6°
-150°	-151.5°	-151.5°	-149.5°
-120°	-121.3°	-122.3°	-120.4°
-90°	-90.1°	-90.1°	-91.3°
-60°	-61.9°	-61.9°	-60.2°
-30°	-30.7°	-30.7°	-30.1°
0°	0.5°	-0.5°	-3°
30°	29.8°	29.5°	29.1°
60°	60.3°	60.3°	60.2°
90°	90.6°	89.6°	90.3°
120°	119.9°	118.9°	120.4°
150°	148.2°	148.2°	150.5°
180°	179.5°	179.5°	178.6°

7 Conclusion

This study developed a motion capture system for multi-sensor real-time motion tracking and implemented an HMC system for the experiment. By comparing the actual measured angles to the content angles, this study verified the effectiveness of the real-time motion tracking technology. As future work, readings from accelerometers will be used to express sportic human motion more effectively, and in-depth studies will be performed on user recognition models.

Fig. 13 Use of HMC system

References:

[1] Poong-Woo Jeon and Seul Jung, "Teleoperated Control of a Mobile Robot Using an Exoskeleton-Type Motion Capturing Device Through Wireless Communication", Journal of Institute of Control, Robotics and Systems, Vol. 10, pp. 434-441, 2004.

[2] Changho Han, Soonchul Kim, Pyoungwon Seo, Youngkee Ryu, and Choonsuk Oh, "A Study of Improving the Accuracy in the 3D Data Acquisition of Motion Capture System", Korea Computer Congress, Vol. 35, pp. 503-506, 2008.

[3] Chil-Woo Lee and Hyun-Ju Lee, “Gesture Recognition and Applications”, ICT-based Image, Arts, Contents and Technology Workshop, pp. 1-11, 2001.

[4] Pyeong-Won Seo, Yu-Geon Kim, Chang-Ho Han, Young-Kee Ryu, and Choonsuk Oh, "Development of a Low-cost Monocular PSD Motion Capture System with Two Active Markers at Fixed Distance", Journal of the Institute of Electronics and Information Engineers, Vol. 46, pp. 61-71, 2009.

[5] Jihong Lee and Insoo Ha, "Multi-Sensor Motion Capture System Using Accelerometers and Gyro Sensors“, Proceedings of the 14, KACC, Oct 1999

[6] Yunyoung Nam, Yoo-Joo Choi, and We-Duke Cho, “Human Activity Recognition using an Image Sensor and a 3-axis Accelerometer Sensor”, Journal of Internet Computing and Services (Vol. 11, No. 1).

[7] Jijang Jeon, “Structure of Kinect in Extraction of Three-dimensional Information,” Electronics Technology Vol. 26, No. 2, Consecutive No. 301, pp. 31-42, 2013.02.01.

[8] MicroStrain,Inc, “3DM-GX3®-15, 3DM-GX3®-25 MIP Data Communications Protocol”,Firmware Versions 2.0.00 and higher, DCP Manual 8500-0024 Revision 003, June 27, 2012.

[9] Nokia, “OpenGL Tutorial Release 1.0”, Qt Learning, June 01, 2012.