3dID: a Low-power, Low-cost Hand Motion Capture Device

Michele Sama, Vincenzo Pacella, Elisabetta Farella, Luca Benini, Bruno Riccó
DEIS University of Bologna
v.le Risorgimento 2, 40136 BO Italy
$\{msama, vpacella, efarella, lbenini, bricco\}$ $@deis.unibo.it$

Abstract

This paper presents a novel input device design for capturing gestures. The system is based on commodity components and combines accelerometers, gyroscopes and bend sensors. It is a low-power, low-cost hand device, characterized by extreme wearability thanks to wireless communication support and small form-factor. It can be used as a stand-alone platform or combined with other wireless sensor nodes in a body area network. The system has been tested as input interface for moving a virtual three-dimensional hand in real-time.

1. Introduction

The tremendous advances in wireless connectivity, interface tools design and embedded system deployment bring sophisticated intelligent technologies into people's ordinary life at home, workplace and in other common private and public spaces. Interaction methods deeply affect the usefulness of being surrounded by a responsive virtual/physical world made of embedded technology. Gesture interfaces contribute to the feeling of immersion and effectiveness of interaction. At present, high-end commercial solutions are targeted for highly-accurate motion capture for realistic animation of synthetic characters. Such systems are cumbersome fixed infrastructures, expensive, needing frequent and complex calibration processes. Therefore, they need specialized operators and limit user mobility to a restricted area. In this work we explore natural interface domain in the particular research field of gestures and hand motion capture. Design of a wireless low-cost/low-power solution based on commodity components, called 3dID (Tri-dimensional Interaction Device), is presented. The system is a glove based on a wireless unit [4] implemented with off-the-shelf components chosen having in mind low-power, low-cost constraint together with wearability and usability. At present 3dID is equipped with three different kind of sensors: five bend sensors, one digital three-axial accelerometer and three monoaxial gyroscopes. Components were chosen to achieve compact size and, as a consequence, extreme wearability. Being based upon wireless sensing unit, the device here proposed overcomes other solutions where sensors are connected through cables. The system includes a second wireless units, called gateway, also designed complying with wearability, low-cost, low-power and small form factor constraints. The gateway enables the glove to be interfaced with general purpose systems through a range of wired and wireless standard interface choices (e.g. USB, Bluetooth, RS232, Ethernet). Communication among the gateway and the glove is supported through physical and MAC layer. Moreover, the system we propose can act both as a stand-alone device or as part of a Body Area Network (BAN). As a consequence, the system presents high flexibility, enabling connection of more than one 3dID device, but also of other kind of wireless and wearable units capturing data from different parts of the body. To test our device, a 3D virtual hand simulation environment has been developed.

The reminder of this paper is organized as follows. First we report on related work, then we provide a detailed description of the hardware design. Afterwards, we demonstrate the use of the 3dID as a motion capture recognition system in a vertically integrated application, where an hand avatar is controlled by our glove.

2. Related work

Gesture recognition and motion tracking is an active research field due to the interest coming from many different field of applications: robotics, Virtual Reality, healthcare systems, tele-presence, sign-language translation and recently game applications and motion capture for animation of three dimensional synthetic characters.

Vision systems for capturing movements have been largely explored. They are highly accurate but expensive systems for absolute positioning in space. Optical tracking need a line-of-sight between the camera and the tracked object and it is usually influenced by environmental lumi-

3-9810801-0-6/DATE06 © 2006 EDAA

nance. In addition, it is difficult to use it for real-time applications since the system needs post-processing of the data to extract motion information. Moreover, a general weakness of this approach is the need for frequent complex calibration processes, decreasing usability for the average user. Alternative solutions are electromagnetic tracking devices, less expensive w.r.t. vision based systems, but still not adequate for the consumer market and suffering from magnetic interference problems from metals objects and difficult calibration.

Relatively less accurate, but successful both in research and commercial field is use of low-end systems for hand motion tracking [15]. Gloves are more cost-affordable devices compared to high-end solutions, they are mainly based on use of off-the-shelf components and sensor fusion. Typical sensors used are inertial and bend sensors. Yet in this area high variability of wearability, usability and costs can be observed. CyberGlove [11] is an interesting device, equipped with a rich set of bend sensors, monitoring different kind of finger movements, but it still costs in a range between 10000 and 15000 USD. As other devices of this kind [12], it is a wired solution, interfacing with general purpose systems through cables and cumbersome additional boxes. In general, the majority of commercial devices only handles fingers bending or hand rotation. The 5DT Data Glove [8], has similar price and it is sold in different configurations. It can be bought with its specific wireless kit, but does not analyse translations. The basic configuration offers only a subset of functionalities, while the more equipped one can manage abduction between fingers but still does not handle translations. The P5 Glove [10] is the only device among those we reviewed with a comparable price but it does not have wireless capabilities. This device needs a desktop mounted receiver for motion capture, and receiver range is limited to 3-4 foot range.

Existing glove devices are often limited by cables or by cumbersome bridge devices to interface with a desktop PC. Our solution is totally on board and is lightweight and small sized. Thanks to the low power optimisation at different design level (e.g. hardware components choice, power optimized protocol implementation) the 3dID, in contrast with other commercial gloves, is completely wireless, equipped with portable small batteries for wearability and mobility support.

Gloves and usage of sensors (accelerometers, gyroscopes and bend sensors) for building interfaces and designing interaction devices is not only interesting for commercial use but it has been largely explored in many research studies since a couple of decades [16]. Perng et al. [14] present a system interface dedicated to hand movement recognition to enable mouse-like input using an accelerometer glove. The glove is equipped with six 2-axis accelerometers on the finger tips and back of the hand. The glove has also an RF transmitter to send data to a personal computer, thus acting as a wireless input device. Using the Acceleration Sensing Glove (ASG) twenty-eight static hand gestures are recognized. The recognition is based on the general use of accelerometers as inclinometers for steady-state measurements, exploiting trigonometrical relationship [13]. Hernandez et al. [5] present a portable glove, the AcceleGlove, based only on accelerometers. The system is used to recognize basic gestures of the American Sign Language. 26 hand shapes are recognized through a three-level hierarchical classifier with high-accuracy.

One major problem faced in many interaction interfaces based on inertial sensors (e.g. [1]) is the compensation for inaccuracies in position and orientation detection due to the integration process. This problem is common to most of the accelerometer based solutions we analyzed. Different approaches have been evaluated to face it, ranging from Kalman filtering [1] to compensation techniques exploiting external measurements or introducing re-initialization, as it happens in [1] where the Zero-Velocity compensation technique is exploited. This technique is not tailored for example to a context of continuous real-time interaction as it could be in virtual environment [2].

3. Hardware Architecture

There are two hardware units involved in this design (Figure 2): a wireless node, which is directly attached to the glove, and a gateway node, whose role is bridging the connection between the wireless node and a generic host RS232, USB, Ethernet or Bluetooth compatible. The wireless node and the gateway share the same modular and stacked design approach, which allows to define hardware layers for each specific task: Power Supply Layer (PSL), MicroController Layer (MCL), Wireless Transmission Layer (WTL), Host Interface Layer (HIL) (see Figure 1); each hardware layer is an independent printed circuit board (PCB) with a common footprint for the cross-layer connectors. In order to obtain a stabilized voltage at 3.3V for the whole wireless node, the PSL is provided with a charge pump integrated device. The MCL is equipped with an Atmel ATMega 8 microcontroller, which provides all the necessary I/O for the sensors and the transceiver. On the wireless node the MCL is also provided with all the necessary pull-up resistors and decoupling capacitors needed to connect the analog sensors. The WTL is based on the TR1001 transceiver, which operates in the 868 MHz European free bandwidth, and a surface mounted antenna. The HIL on the gateway node provides a RS232 interface, a Bluetooth module manufactured by SmartM (called Bluetopaz) and a USB interface based on the FTDI FT245BM. Recently a Ethernet interface has been added to the HIL.

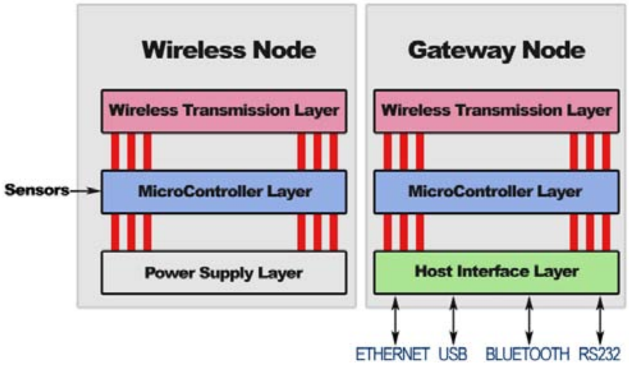

Figure 1. Hardware architecture

The modular design enables different choice for interfacing with the 3dID, enhancing flexibility.

3.1. Bend Sensors

The fingers flexion is measured by five resistive bend sensors (one for each finger) hooked up to five of the eight available ADCs of the microcontroller. These sensors are composed of tiny patches of carbon that change resistance values when bent from convex to concave shapes; average resistance values for such devices are comprised in the range $[25 \div 350]\ K\Omega$. Each sensor is used in conjunction with a voltage divider to provide changing voltages, read by the microcontroller on its analog inputs.

3.2. Gyroscopes

The remaining three ADC ports on the microcontroller are hooked up to three ADXRS300 piezo-gyroscopes (Analog Devices). These sensors measure the angular rate on each axis (X Y Z) and are arranged in an orthogonal configuration. The signal measured is voltage-proportional to angular rate about the axis normal to the top surface of each IC package.

3.3. Accelerometers

The LIS3L02DQ (ST Microelectronics) is a MEMS triaxis digital linear accelerometer with a SPI serial interface. This device is hooked up to MISO and MOSI ports of the microcontroller and provides a 12bit value for each axis.

3.4. Firmware and packet structure

The gateway and the wireless node have different firmware. The firmware on the gateway performs data and command packets forwarding between the host and the node(s), while the wireless node firmware performs several tasks such as sampling from the analog inputs (Bend Sensors and Gyroscopes), reading data from the triaxial accelerometer using the SPI protocol, sending data packet to the gateway, executing commands received from the host/gateway (e.g. START_SENDING, STOP_SENDING,...). The data packet structure illustrated in Table 1 has been designed for a general sensor node, not specifically for the glove. The command packet differs from data packet because it does not have a PACKET_COUNT byte and the data bytes are replaced by a COMMAND_ID byte.

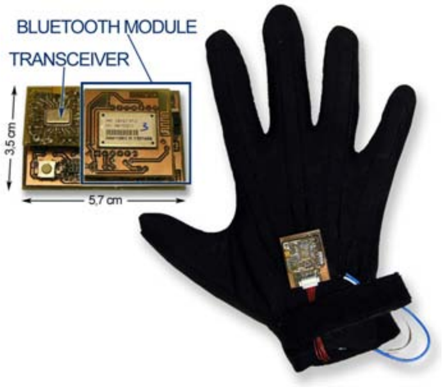

Figure 2. The 3dID glove and the gateway

Byte	Description
START_BYTE	Start byte 'i'
DEVICE_ID	Device identifier
DEVICE_TYPE	Device type identifier
DATA_BYTE[0]	First data byte
...	
DATA_BYTE[N]	Last data byte
PACKET_COUNT	0-255 packet count
END_BYTE	End byte 'f'

Table 1. Data packet structure

The pair [DEVICE_TYPE:DEVICE_ID] acts like a MAC address, identifying the node in the whole BAN. The device identifier is also used to determine the data packet size (each node type of the BAN has a different data packet size, depending on the sensors it is provided with). For the glove, the data packet size is 19 bytes (as explained in details in Table 2)

Sensor type	N of Sensors	N of bytes	Total bytes
Accelerometers	3	2	6
Gyroscopes	3	1	3
Bendsensors	5	1	5

Table 2. Sensors and packet size

4. Software Architecture

Real time response is one of the primary targets in 3dID design, as a consequence, to demonstrate this capability we first implemented a software application in the OpenGL environment where a virtual tridimensional (3D) hand is displayed, directly controlled by the glove. The choice to build a 3D application is driven by several considerations: the glove offers a digital sampling of a human hand so the more logical way to test the consistency of the data stream is to give it the same meaning that it has in reality.

3D applications usually need complex processing and the data input stream must be constant and synchronized to avoid delays and side effects. Thus, by testing the glove with an OpenGL environment, both the real time and synchronization requirements are guaranteed and the data management is not excessively demanding in terms of computation. As a final consideration, it is important to underline that the glove and the 3D application have been developed simultaneously. Therefore, the hand avatar represents a strategic testing tool for the hardware device.

The OpenGL application is composed by two main blocks (Figure 3): the *presentation layer* and the *glove driver*, described in the following subsection.

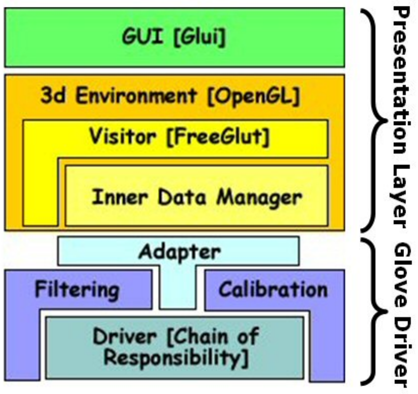

Figure 3. Software components

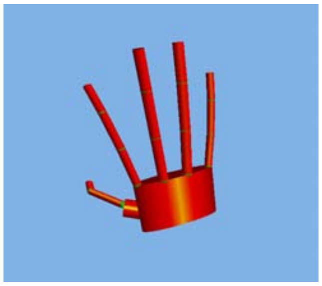

Figure 4. Hand 3D avatar

4.1. The presentation layer

The presentation layer is composed by the 3D environment and the GUI. The environment is the primary component, it contains the OpenGL world environment and the hand avatar (Figure 4). The avatar is the main element of the application; it should represent a human hand so it has a defined shape and specific movements constraints. Following this guidelines applying the *Visitor* pattern (see [9] for Design Patterns) we split the avatar into two subcomponents, the *inner hand data manager* for sensor data management and a runtime selected *Visitor* invoked by the OpenGL main loop that redraws the avatar in the 3D environment. The glove uses five sensors to measure fingers flexion from which we can calculate one bending angle for each finger. In the avatar each finger is composed by three joint and three phalanx, as a consequence one angle for each is needed. The problem has been solved by assigning the angle to each phalanx referred to the previous one: in presence of an input angle x the first node rotation is x, the second is 2*x and the third is 3*x.

The avatar, being a 3D object, needs three spatial coordinates for the position in the 3D environment and three angles relative to the 3D spatial axes for the orientation, therefore the sensor input must be converted into a more suitable value for the driver subcomponent. The avatar visualization is granted by the *Visitor* component. The *Visitor* automatically redraws the 3D scene when the *inner data* are changed. There are two different implementations of the *Visitor*, one is a light and schematic object that draws a wireframe skeleton, the second, more complex than the previous, is based upon the open source library FreeGlut [6], a modern extension to the GLUT OpenGL library. This one has been selected as the final *Visitor* because a third implementation with Nurbs 3D surfaces appears to be computationally demanding for real time redrawing purpose. Actually no frame limits or frame rate controls are implemented and the redraw speed is the slower between the OpenGL main loop speed and the driver data-load speed.

During development and testing phases of both the hardware and software systems the need has emerged for direct control of the *inner data component* (e.g. for debugging purpose of the software and for individual testing of glove sub-systems). Due to this need, a control GUI has been developed with the GLUI open source library [7]. The main task of this user interface is to enable or disable the data stream, starting or stopping the driver component. Alternatively data can be simulated using graphical controls. This capacity has been crucial in order to test and use the glove functionalities in the starting phase of the design and implementation and it is at present very useful for testing each sensor separately.

4.2. The Driver Component

The driver component is the bridge between the glove and the software systems and it is application independent. Its main purpose is to read data from the glove connected communication port and to convert them into a *Packet* object, which is pushed by an *Adapter* into the *inner hand data manager* subcomponent. The design of the low-level component of the driver, which runs as a stand alone thread, is based on a pattern called *Chain of Responsibility*. Each *Handler* of the chain is able to read from the input data stream and is able to check the consistency of read data. If the packet is marked as non-valid, the chain is restarted. This implementation grants two different properties: a real time error discovery and filtering and a lightweight refactoring in case of changes in the input stream protocol (see also section 3.4). Data received can be sensor data or synchronization data. Any value between 0 and 2^8 is acceptable in the sensor data field. Otherwise, in the synchronization field the value is fixed and is predictable. Each reading *Handler* of the chain is able to check the validity of their managed byte or to reset the reading chain. The chain starts with the *EndByteReader*, which seeks for the ending byte of a packet, then the chain passes the control to the *StartByteReader*, which reads the first byte in the stream and if it is different from the start byte, it restarts the chain. A similar approach is used for every control byte of the protocol such as the DEVICE_ID byte, the DEVICE_TYPE byte and the PACKET_COUNT byte. This modular design enables easy protocol update.

Once a valid packet from the input is stored the driver must do some correction into the read value converting the read data into a 5+3+3 float structure: five floats for finger flexion, three for 3D space position and the last three for 3D orientation (yaw, pitch and roll). In preliminary tests we noticed that a direct use of the 16 bits of accelerometers value is affected by noise due to the intrinsic nature of the hardware component, to human uncontrolled movement and to ambient imperceptible vibrations that will result with an on-screen avatar tremor. Consequently, a configurable bitmask based filter has been inserted into the driver. The bend sensors must be calibrated because their value range is different from one to another and could be ageing and temperature dependent, therefore we perform a software conversion by reading the minimum and maximum value and normalizing read data into this range.

The GUI offers a control button for starting/stopping the calibration routine; while it is running, the user who wears the glove should try to move each finger as more as he can in order to provide a more accurate reading.

A critical task, at present under development, is rotation and translation handling. This is one of the main driver functions, but its development is based on gyroscopes and magnetometers complete integration that is not fully achieved at the moment. Movements results from the combination of rotation and translation. Accelerometers in steady-state condition can accurately measure rotation with respect to axes non perpendicular to the gravity [3]. Considering the 3D space as a sphere and subdividing it into eight parts in a static context accelerometers and magnetometers values are sufficient to identify the object orientation. Furthermore, in dynamic conditions rotations are extracted from the combination of gyroscopes and accelerometers data.

5. System Evaluation

5.1. Power Consumption Analysis, Costs and Form Factor

The power consumption on the wireless node is divided as follows:

COMPONENT	Avarage Power Consumption (mW)
Power stabilizer	0.2
Microcontroller, Accelerometer, Bend Sensor	27.72
Transceiver	25.08
Gyroscopes	60
Total	113

Table 3. Power consumption of single components

Considering values in Table 3, at 50HZ acquisition rate and 9600 bps of throughput, the 3,3 V/500 mAH battery provided with the wireless node has an estimated average duration of 14 hours. The prototype device (glove and gateway) costs around 150 Euro in the configuration equipped with Bluetooth capability.

5.2. Wireless Performance Analysis

To evaluate the system we tested the performance of the 3dID in terms of packet loss. We measured how the distance between the gateway and the glove affects the correct reception of packets. The test consisted in counting packets correctly received and comparing the value obtained with PACKET_COUNT. Inconsistencies between the two values provided an indication of packet loss. Table 4 shows that almost all the packets are received correctly in a distance range of 5m. In the range between $[5 \div 8]$m the system decreases its performance but it still usable. For distances greater than 8m the system packet loss exceeds 40%. Thus, 3dID usability is guaranteed within 5m, which is suitable for typical usage of such device.

Distance (m)	Packet Loss (%)
0,05	0
1	0
5	4,83
8	42,6

Table 4. Percentage of packet loss

As a conclusive observation, we verified that the system is affected by interference coming from cellular phones. Thus we plan to substitute the transceiver, actually working at 868Mhz, with one working at 2,4Ghz.

6. Conclusions

We described design and implementation of a wireless low-power low-cost wearable glove for natural interaction through gestures. The hardware device is equipped with wireless networking support, hardware-software interface to a general purpose system and a graphical 3D environment has been developed for debugging and real-time testing of the glove. On-going work is aimed at implementing a quaternion based orientation filter. A similar approach is used in airplane inertial guidance systems. To the purpose we will add magnetometers and implement a Kalman filter at driver level.

References

[1] W.-C. Bang, W. Chang, K.-H. Kang, E.-S. Choi, A. Potanin, and D.-Y. Kim. Self-contained spatial input device for wearable computers. In *ISWC*, pages 26–34, 2003.

[2] R. Barbieri, E. Farella, A. Acquaviva, L. Benini, and B. Riccò. A low-power motion capture system with integrated accelerometers. In *CCNC*, pages 418–423, 2004.

[3] E. Farella, A. Acquaviva, L. Benini, and B. Riccò. A wearable gesture recognition system for natural navigation interfaces. In *EUROMEDIA International Conference*, pages 110–115. Eurosis, 2005.

[4] E. Farella, A. Pieracci, D. Brunelli, A. Acquaviva, L. Benini, and B. Riccò. Design and implementation of wimoca node for a body area wireless sensor network. In *IEEE International Conference on Sensor Networks (SENET 2005)*, pages 342–347. IEEE Press, 2005.

[5] J. L. Hernandez-Rebollar, R. W. Lindeman, and N. Kyriakopoulos. A multi-class pattern recognition system for practical finger spelling translation. In *ICMI*, pages 185–190, 2002.

[6] http://freeglut.sourceforge.net/.

[7] http://sourceforge.net/projects/glui/.

[8] http://www.5dt.com/products/pdataglovemri.html.

[9] http://www.dofactory.com/Patterns/Patterns.aspx.

[10] http://www.essentialreality.com/.

[11] http://www.immersion.com/3d/products/cyber_glove.php.

[12] http://www.vrealities.com/glove.html.

[13] A. Paiva, R. Prada, R. Chaves, M. Vala, A. Bullock, G. Andersson, and K. Hook. Towards tangibility in gameplay: building a tangible affective interface for a computer game. In *ICMI '03: Proceedings of the 5th international conference on Multimodal interfaces*, pages 60–67, New York, NY, USA, 2003. ACM Press.

[14] J. K. Perng, B. D. Fisher, S. Hollar, and K. S. J. Pister. Acceleration sensing glove (asg). In *ISWC*, pages 178–180, 1999.

[15] D. J. Sturman and D. Zeltzer. A survey of glove-based input. *IEEE Comput. Graph. Appl.*, 14(1):30–39, 1994.

[16] T. G. Zimmerman, J. Lanier, C. Blanchard, S. Bryson, and Y. Harvill. A hand gesture interface device. In *CHI '87: Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*, pages 189–192, New York, NY, USA, 1987. ACM Press.