Ergonomic Low Cost Motion Capture for every day health exercise.

Dennis Majoe\textsuperscript{1}, Michel Estermann\textsuperscript{1}, Nicola Ranieri\textsuperscript{1}, Jürg Gutknecht\textsuperscript{1}
\textsuperscript{1}Computer Systems Institute, ETH Zurich, Switzerland
dennis.majoe@inf.ethz.ch

Abstract

Pervasive applications often require context information about the posture, ambulatory and work activity of the user. This paper reports ongoing work on the development of a low cost ergonomic approach to motion capture that provides the opportunity to examine diverse body movement in home and office environments. In this research case it is applied to the automatic training of Tai Ji Quan at different levels of complexity. Results of the sensor design, communications architecture and human body computer motion modeling are presented.

Keywords: Wearable computing, motion capture, body area networks, wireless sensors.

1. Introduction

Research in the Computer Systems Institute, ETH Zurich, is directed at Health and Wellness through wearable computer based assistive devices. Rehabilitation and fitness will often involve the physical side of human activity. Our experience in this domain shows clearly that pervasive applications that require motion sensors to detect the physical posture and displacements of the users limbs and torso have some very specific and some very general requirements.

Specifically the motion can be detected using a wide range of intelligent sensors, leading to design choices for the technology to be applied. A wide variety of motion capture systems employ cameras and or infra red sensors mainly to measure absolute position and translation information. [1] Alternatively several systems measure purely rotational information for each limb and use forward kinematics to calculate the absolute position of all limbs. These systems are predominantly used in the animation and multimedia industry and the cost of the equipment is in almost all cases outside the scope of the day to day end user. A review has shown that lower cost systems are bulky, wired and as a result non ergonomic.

A general system was required to bring wearable motion capture to the end user so that a new range of applications could be realized. This required a system that provided numerous sensors for each limb, ergonomically wireless, very small and unobtrusive, with long battery life and limited calibration requirements. In addition the final system should be within reach of the finances of typical computer games user.

Finally the sensor system should be applied generally in an Internet friendly architecture in which motion data streaming is easily linked to 3D worlds and avatar modeling.

2. Motion Capture Design Choices

In the wearable sensor context, there are several design choices to be made to resolve the optimal motion capture system.

Primarily the user's context, their environmental conditions and physical movements must be adequately measured using suitable sensors that achieve the necessary accuracy, ergonomics, data rates, reliability and so on.

Almost all motion capture systems, known to the authors, are subject to errors and limitations. Ambient factors such as electrical noise and where the measurements are conducted will influence an outcome. For example some systems require a specific lighting or limit the user to move in a restricted domain.

Often these systems must be used in constrained conditions for them to work reliably and some type of calibration is needed. Certain systems use data from more than one type of sensor and use data fusion to obtain higher levels of accuracy, reliability and noise hardening. [2] The benefit of data fusion of multiple sensors must be balanced against the difficulty to ergonomically implement multiple sensors in a small package in real hardware. Battery life is another key point when hardware is being designed since the system, which must have very small batteries and wireless consuming energy on the RF side as well as the sensor side.

The approach taken in this research has been to put more emphasis on a design with low financial sensor cost, and high user friendliness. Although this may mean motion accuracy is limited, it is believed that software will process this data in such an intelligent way so as to achieve the fundamental needs of the application domain.

2.1 The Rotator Sensor System

The motion capture system developed was called Rotator since it relies on measuring the angles of
rotation of each limb. It has the following characteristics.

The motion sensor measures the 3 degrees of rotation about the 3 axis of the sensor. One sensor is attached to one limb. The sensor measures earth’s gravity acceleration vector and the earth’s magnetic field vector in order to compute the rotations.

Acceleration measurement of earth’s gravity is done using standard MEMS based accelerometer devices. Previous work [3] had shown that the acceleration forces that the human body generates are most often insignificant compared to the gravity force. When the forces are high then this occurs for very short periods. Therefore in most cases the pitch and roll angles of a limb relative to the vertical axis adequately meets our motion capture requirements and the vertical reference is guaranteed for most environments. However rotations about vertical cannot be measured with the gravity vector alone. Therefore the earth’s magnetic field is used as the secondary sensor data. The two sources of measurement data are fused to provide the final rotation values. This is done by calculating the pitch and roll angles using the gravity vector alone and then adding the magnetic azimuth signal, compensated for pitch and roll using well known compensation methods. [4]

Although it is weak, the earth’s magnetic field is generally sufficient to provide the azimuth rotation signal. Ferromagnetic materials near the sensor could result in localized angular rotation offset that can be easily compensated for provided all sensors are similarly affected. In practice the main problem occurs when there are field disturbances due to moving metal objects near individual sensors. In this case the system would record uneven and untrue variations in the limb’s azimuth measurement. This sensor system is therefore limited to use outdoors, in office and home spaces but should not be used in close proximity to large moving metal structures.

Calibration is a simple process requiring each sensor to be rotated to the maximum and minimum sensed signals. The maximum and minimum values are used to derive offsets and scaling terms. 

2.1 Radio Data Transmission BAN

The Body Area network applied in this system is based on the Nordic nRF24E1 microcontroller. This provides a radio transceiver for the world wide 2.4 - 2.5 GHz ISM band and was chosen in particular for its low power consumption with an RF current consumption of only 10.5 mA in TX mode (output power -5dBm) and 18 mA in RX mode. The transceiver features a data rate up to 1 Mbps, multichannel operation with 125 channels.

The basic architecture of the Body Area network (BAN) requires a single master node and multiple sensor slave nodes. Each sensor slave sends RF borne data to the master for collection and transfer to the host computer, a PC, which runs the main application. The BAN may consist of up to 255 sensors sharing the 1Mbps bandwidth.

In order to improve the RF transceiver data handling, a Frequency Hopping algorithm has been implemented in which the master and all slaves follow a specific random walk based channel selection algorithm. By randomly using different channels the RF link is more robust and less likely to be compromised by other devices using the same frequency channel.

2.3 Data handling Protocol

Today there exist several standards for wireless connecting devices, such as IEEE 802.11 (WiFi), IEEE 802.15.1 (Bluetooth) and IEEE 802.15.4 (ZigBee). IEEE 802.11 is not intended for devices with power consumption constraints as in sensor networks and rarely used for small battery powered sensor nodes. Sensor network implementations are often based on the IEEE 802.15 standards however both Bluetooth and ZigBee have major constraints in terms of supporting topologies with multiple sensors.

The limitation of seven slaves in a Bluetooth piconet and the interference between multiple piconets significantly limit the use of Bluetooth for motion capture sensor networks. For ZigBee the limited bandwidth and MAC protocol design also restrict the use of the technology for body sensors. Further, both of these communication stacks are complex and have a rather large footprint. It is difficult to implement them on a 8-bit micro-controller with limited memory without making major modifications.

Additionally unlike other applications such as environmental monitoring using multiple sensors, motion capture requires a combination of large sensor count, low power consumption and very high bandwidth.

The MAC protocols can be distinguished into TDMA and contention-based protocols. Contention based protocols like S-MAC and T-MAC use coordinated sleeping and low-duty-cycles to address the major sources of energy waste, i.e. overhearing, idle listening and over emitting. Other examples of contention-based MAC protocols are B-MAC, SEESAW and WiseMAC.

TDMA-based protocols first establish a schedule where each node is assigned a timeslot (or multiple timeslots). The nodes are then able to communicate in the assigned timeslot without provoking collisions and do not suffer from overhearing. Pure TDMA protocols require global time synchronization and network topology changes may pose problems. Many proposed protocols for sensor networks use therefore a combination of the contention-based and TDMA-based approach.
Time Division Multiple Access (TDMA) seems to be the best and only feasible approach to meet the real-time requirements and reach the highest possible throughput. Demand-based, e.g. Polling, and contention-based approaches require too much time for necessary message exchanges and may result in high latency.

The protocol proposed here is similar to the protocol described by Aylward et. al [5] but introduces additional features to meet the slightly different application requirements, namely to increase the reliability for some application scenarios and allows to increase the number of possible sensor by using lower sampling frequencies for some nodes. Further a contention-based setup protocol is proposed to assign available timeslots to nodes at anytime, allowing a scalable network.

2.4 Wireless Protocol Overview

To avoid time consuming radio reconfiguration, the network nodes use the DuoCeiverTM ability of the nRF2401 radio subsystem where possible. Three frequency channels are used; a Transmission Channel, a BAN Broadcast Channel and a Common Broadcast channel.

All nodes share the same Common Broadcast Address assigned to the Common Broadcast Channel. This channel is used by the master to discover all sensor nodes in the network during the initial Setup Phase. By using an address known to all nodes we ensure that sensor nodes can join a network.

After discovering nodes the master assigns a unique BAN Broadcast Address to all associated sensor nodes in the network. This ensures that sensor nodes do not react falsely to broadcast messages received from another nearby network master.

Finally a unique Transmission Address for the Transmission Channel is assigned to every sensor. The Transmission Address is formed from the three most significant bytes of the BAN Broadcast Address plus the Node ID. A dedicated transmission channel allows us to have different payload lengths for broadcast, data and data acknowledgment packets and the time to reconfigure a channel (payload length and address) can be saved. The BAN Broadcast Channel is always separated by 8MHz from the Transmission Channel.

2.5 Sensor configuration and sampling

Using the Common Broadcast Channel any sensor may request to be accepted by the master to join the network. This is achieved with a simple contention based algorithm in the case two or more sensors attempt to do so at the same time. Once accepted the master assigns a unique address to the sensor.

In order to begin sensor sampling a communication cycle and the TDMA schema is triggered and synchronized by a SAMPLE message from the master node to all sensor nodes on the BAN Broadcast Channel. DATA is then sent from the nodes to the master node during the assigned timeslot using the Transmission Channel. Depending on the number of nodes and required throughput the DATA packets will be acknowledged or not.

![Figure 2.1](image1.png)

Figure 2.1 shows a single TDMA sensor data transmission cycle for a sampling frequency of 90Hz without acknowledgment for high throughput (upper) and with acknowledgment and retransmission for more reliable communication (lower).

It may be desirable to have sensors with lower sampling rate coexisting with sensors of higher sampling rate. In this case multiple sensors can be assigned to the same timeslot using the timeslot interleaved, e.g. every second transmission cycle only. With this the maximum number of possible sensor nodes in a BAN can be increased. Figure 2.2 shows this principle.

![Figure 2.2](image2.png)

2.6 Battery Power

The main power consumption in the sensor comes from the RF section and specifically the Receiver. Since the receiver can be switched on and off, it is absolutely necessary to only enable it when the sensor is likely to receive the sample command from the master. Having received it, the sensor can switch off the receiver and thereby save power. The TDM approach conveniently suits this power saving strategy. This is the basic method to save battery power since the receiver takes 18mA of the total 21mA average power consumption.

2.7 Early PCB Prototypes

The first prototype sensors have been made of similar size as a man’s watch and packaged in a plastic case with easy access to change the CR2032 coin cell.
battery. With the battery the weight is approximately 35 grams. A range of fixing methods are possible however it was found that a simple Velcro arrangement provided a flexible approach to reliably bind the sensors to users limbs using neoprene and fabric bands. As in figure 2.3 one can see the PCB on the left and it being worn on the wrist in its plastic case on the right.

The way the user wears the sensors is important since they must be low profile to avoid being knocked and must stay in the same orientation as they are assumed to be in the bones model.

3. Data Streaming and Processing

Once data is received by a network master it is sent by serial link to the host PC. However it was foreseen that a simple one to one data stream connection approach would be limiting in the case that multiple users were interacting and using motion capture simultaneously in real-time. In addition the users could be at the office or in a sports hall or at home and they would join each other over the Internet.

To ensure that multiple users could coexist and data could be treated by a single application or a set of applications, the data handling was augmented with an architecture comprising Producer Clients, Consumer Clients and a Central Server. Producer clients read data from master nodes, which receive data from a configured set of sensors. Producer clients may read data from the Rotator system or any other real time system. The Producer clients make a TCP/IP connection to the server and deliver data streams to the server following predefined protocols for registration data format and deregistration. The produced data may be accessed by Consumer clients that also make a socket connection to the server using similar predefined registration and data protocols.

An end user may be a single Consumer client such as a passive spectator or may combine a Producer client and interact with motion data. This architecture allows multiple users to interact across the Internet, using either this motion capture system or any other system, provided the data stream can be correctly encapsulated by the Producer and Consumer for acceptance by the server.

3.1 Forward Kinematics

Whilst alternative motion capture systems measure absolute position of markers on the user body, this system’s sensors provide a measurement of each limb rotation relative to the vertical and azimuth. The actual absolute position of the user’s limb extremities, for example, must be estimated by using the rotation data and the length of each limb to calculate each limbs position in 3D space. Forward kinematics is the technique used to do this calculation and the user’s limb lengths can be measured or approximated statistically [8]. The position of every part of the body is calculated starting from a root body limb such as the waist and hips. The rotation data is applied to this body part to find its orientation in space. Typically the root has two or more connecting limbs. In this case there are three connecting the chest and the two legs. Each connection point corresponds to a vector displaced from the centre of the root. These connection points are therefore rotated in space by the same root sensor rotation values and provide the starting point for the next limb.

The next limb is represented as a vector whose origin is one of the connection points on the root. This limb vector is rotated in space by the corresponding sensor angles and the limb end defines the origin for the next limb vector. In the case of the chest it has two following connection points and therefore is made of two vectors.

Therefore a 3D reconstruction such as in Figure 2.4, of the user can be calculated providing clear absolute translational information for each limb extremity e.g. the hands. When a generic bones length model is used for the calculations, every user’s body can be characterized by this model. In this way the movements of different users can be directly compared despite the fact that actual body dimensions between different users may vary significantly.
4. Prototype Sensor Results

The real time motion capture system was tested using 25 sensors. In this test 15 were set to the maximum sample rate and 10 set to half the maximum rate.

4.1 Configuration

Automatic configuration by the master node makes configuring the sensors extremely easy to do. By simply switching on the sensors one by one, each sensor can request assignment. The master assigns a time slot and address and the host computer can monitor the configuration list and provide status information of the state of all sensors assigned. This process has two main benefits. When a sensor battery has to be changed the sensor that is switched off is automatically de-assigned by the master due to the timeout and then reassigned again when the sensor is switched back on.

Since the sensor address is assigned by the master, a sensor may be used in one application and then in another without the worry to check that its address is unique to the application.

4.2 Sample rates

A 90 Hz sample rate across 25 sensors was achieved, with each sensor sending 16 bytes of signal and headers information per sample. Although this was much better than expected, a bottleneck for data is the current implementation of a RS232 communication link between the host computer and master node. This will be addressed by replacing the RS232 serial based master node with a USB equivalent.

The sample rate is related to the required sensor azimuth rotation angular resolution which is currently 2 degrees. To achieve 0.5 degrees angular resolution, the magnetometer needs more time to integrate the sensed signals and as a result it will reduce the calculated sample rate down to around 50 Hz. So clearly a tradeoff must be made between sample rate and azimuth resolution.

4.3 Battery Life and voltage

The sensors can run at this sample rate for 11 hours on a CR2032 coin battery. This seems to outperform other similar implementations of a sensor network. [9] [10]. However the transmission and reception reliability is severely reduced with falling battery voltage. Therefore a reliable system would run for around 4 hours at a range of around 2 meters from the master node. Since the battery voltage is read by the sensor microcontroller, an automatic online linear compensation for the acceleration sensors can be carried out with falling battery voltage. Also an alarm can be raised to change the battery.

4.4 Forward Kinematics and motion noise

The bones model was programmed for a 10 sensor system in which the sensors are attached to the arms legs waist and chest. Using DirectX graphics libraries, the bones model may be expressed in a simple skeleton avatar. The avatar may be controlled in real time by the motion capture system.

It has been found that there is a significant level of accelerometer noise in the data. Some of the noise is due to power line noise, especially as the RF transceiver is switched on and off and the circuitry being inadequately suppressed. Some noise exists on the accelerometers and needs to be filtered using passive analogue circuitry. The noise currently translates to approximately 2 degrees of angular rotation. The effect of noise is to slightly wobble the avatar and it is worse when the accelerometers are near their full range deflection. Software filters are not used at this stage as it is intended to limit any delays between real motion and rendered motion of the avatar. The measured latency is around 20 milliseconds.

5. Related and Future work

The current prototypes are already being used to perform motion capture of Tai Ji Quan movements. In this work the motion data can be used to assess if the Tai Ji student is correctly maintaining a given static position. By looking at the bones model of the teacher’s body and the student’s body rendered in the same 3D space, the student can in real-time position their limbs in exactly the same position as the teacher. Since the teacher can be remote and since many students can in principle act as Consumers of the teachers motion data, many students can be trained simultaneously by one teacher. The teacher need not address each student since the student self corrects their position. This has proven effective in ongoing trials of the system as depicted in Figure 5.1.

![Figure 5.1](image-url)
5.1 Future sensor modifications

The main areas for immediate attention are the removal of the RS232 bottle neck. This will be achieved by using a USB based alternative to the master node implementation. With a 1Mbps transfer speed the full sensor system will be useable at full data rates.

Calibration of the sensors is currently a manual process that requires automation. The magnetic calibration will be performed by using a mathematical model for the 3D locus of the magnetic vector, whose parameters will be estimated on a continuous basis using a learning algorithm such as recursive least squares. The accelerometer calibration will be simplified by providing square edges on the sensor package to allow the user to simply position the sensor in each of the six possible standing positions and the software will take a snapshot of the maximum and minimum signal values.

6. Conclusions

An ergonomic motion capture system has been developed which can provide a real-time measurement of the body position using a simple bones model as applied in avatar animation.

The system has addressed the specific problems of battery power economy as well as an easily configurable and high data rate body area network for wireless sensors.

While the final cost of production level sensors cannot be accurately judged at this stage, it is believed that a 16 sensor system, which should provide measurements for most body positions, would cost an order of magnitude less than current commercial systems offering the same level of real-time and accuracy.

A specific architecture has been developed and used to show that the sensor data may be transmitted over the internet in a manner supporting a variety of combinations of one to one, many to one and many to many interactions in real time.

The system has been used to progress the development of a Tai Ji Quan trainer application and has been demonstrated working in real time at the recent Shanghai Science Festival at the Pudong Expo Science Exhibition Centre in Shanghai. [11]

Acknowledgement

Acknowledgements go to MA Systems and Control Limited, U.K., for their contribution to the electronics design and assembly of the sensor system.

References


