Changing Interfaces Using Natural Arm Posture – A New Interaction Paradigm for Pedestrian Navigation Systems on Mobile Devices

Ceren Kayalar and Selim Balcisoy

Sabanci University
Computer Graphics Laboratory
Istanbul, 34956 Turkey
`ckayalar@su.sabanciuniv.edu, balcisoy@sabanciuniv.edu`

Abstract. This paper presents a new interaction technique, which is based on arm posture recognition, for mobile computing devices to switch between different visualization modes seamlessly. We implemented a pedestrian navigation system on Pocket PC, which is connected to a GPS receiver and an inertial orientation tracker. In the global coordinate system, user's position is tracked with GPS data, and in the local coordinate system user's arm posture is mapped into two application dependent states with inertial orientation tracker data. Hence, natural interaction and different levels of information is provided by processing orientation tracker data. As unnecessary computation and rendering increase power consumption in small devices, we introduced another state to our system, which saves battery according to the user's idle arm posture.

Keywords: Context-aware retrieval, mobile computing, interaction techniques, posture recognition.

1 Introduction

Location based services are becoming more available for different use cases, most notably navigation and information services. Vehicle based navigation systems became commercially successful systems as they can provide accurate data consistently. Most car navigation systems combine GPS and car telemetry data with Geographical Information Systems (GIS) data. One important advantage of vehicle systems is the available infrastructure to access the telemetry and other electronic systems (radio, lights, phone) and estimate the user context and adjust the system accordingly by adjusting the visual and audio characteristics, such as loudness of the radio, of the system if necessary.

Pedestrian navigation systems also require precise estimation of a user's context, which consists of environmental entities affecting the interaction between a user and an application. Context-aware systems use context to provide relevant services to the user, where relevancy depends on the user's task such as walking around, gazing in a particular direction [1].

With the advent of more capable mobile terminals such as mobile phones and PDAs it has become possible to perform visually appealing 2D and 3D renderings in real-time [15]. These can be maps, task specific applications (rescue, cultural heritage), and interactive guides; and have variable visual complexity. The size constraints of mobile devices make them cumbersome with minimal space for user interfaces. Moreover users expect to retrieve information using mobile computing devices with less effort. Therefore creating natural interaction techniques for mobile devices is an important challenge. One such interaction technique is to utilize users hand posture or gestures to trigger tasks, switch between applications or play games. With the advent of camera phones, first examples of such interfaces begin to be developed [13]. Even the computer vision algorithms are adequate for simple gesture recognition tasks on limited performance terminals; those techniques have not become mainstream, because of their high power consumption.

In this paper we are investigating context estimation methods for PDAs. In case of tour-guide applications we can simplify the context of a mobile user into three: investigation, navigation and idle. In the investigation state user is analyzing the façade of a building or an artifact in detail, where the application should present detailed visual information, if possible, using an Augmented Reality framework [6, 14, 17]. However at the navigation state the application should present 2D or 3D maps of the area with task specific data: street layout, names etc. The application is expected to detect also the idle state conditions when the user is not looking to the screen and trigger power save actions such as CPU clock adjustment. We propose to use a single orientation sensor to estimate the context of the user. This estimation is based on a novel posture recognition technique. The location will be computed using data from GPS sensor.

We first give a review on related work. In section three we describe our recognition method in detail. Afterwards we present a prototype system and case study to illustrate our ideas. The final chapter concludes the paper and presents our ideas on future research.

2   Related Work

Location based services require estimation of mobile users location with adequate accuracy. Depending on application's needs this accuracy can be between kilometers or centimeters. Moreover most of the services require orientation data, at least the heading. GPS is the most prominent technology around for delivering up to centimeter accuracy (Differential-GPS). However, the fact that it only works outdoors is a major drawback. An example GPS based guide application is developed by Burigat and Chittaro [4]. Several alternative positioning technologies such as GSM-Triangulation, Wireless LAN Triangulation can estimate one's position indoors, using access points with well known positions [5, 9]. A recent work of Peternier et al. [12] presents an integrated mobile mixed reality system using Wireless LAN technologies. Another possibility is to track walking data for positioning purposes. There are several research groups working on walking behavior and pedestrian tracking [11, 7].

Until now all location based services we have investigated rely heavily on precise position and orientation of the user in global coordinate frame. However to create natural interfaces for handheld devices we need to recognize upper body postures, most importantly the position and orientation of the terminal relative to user's head.

Recognition of upper body postures and gestures is studied by several research groups: Amft et al reporting a Hidden Markov Models based gesture recognition algorithm up to 94% success rate for eating and drinking arm gestures [2].

Bao and Intille developed a system to detect everyday physical activities using five body-worn wire-free biaxial accelerometers with high accuracy (84%) [3]. They report that even with two sensors – thigh and wrist their recognition performance does not deteriorate. Similarly Lee and Mase report up to 91% success rate with their wearable sensors for activity recognition [10].

Widgor and Balakrishnan implemented a novel text input technique for mobile phones using one 2 DOF tilt sensor (acceleration sensor) to recognize user hand actions [18]. They report that 20 to 50 Hz sampling rates should be required for robust tilt implementation.

3   Posture Recognition

As presented in the previous section, all of the orientation sensor work is based on either to assist precise tracking and positioning of the user in space or gesture recognition using several sensors.

Our goal is to create a stable differentiation mechanism between several hand postures and map them to several application dependent contexts.

The developed recognition algorithm is based on state transitions triggered by time-line analysis of orientation and angular velocity of the sensor. The orientation and angular velocity are the natural arm posture data of a pedestrian, which can be used as context widgets for a context-aware application.

The angle between user's forearm and upper arm is obtained from the orientation sensor as pitch angle, α, and analyzed to recognize different postures. We have gathered sample data from mobile users with various walking speeds, while moving their hands between three postures:

- *vertical*, where pitch angle is around 0°,
- *horizontal*, where pitch angle is around 90°,
- *idle*, where the hand may move freely (Fig. 1).
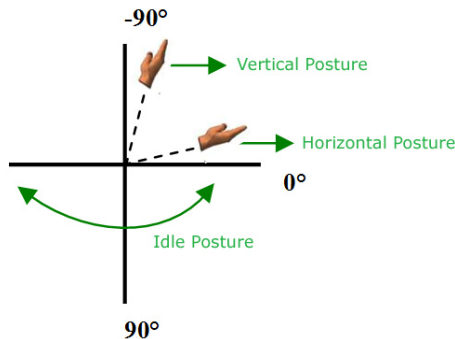
Fig. 1. Drawing of target three hand postures from side view

Fig. 2 shows pitch angle measurements of the user's arm movement in three differ-ent conditions: standing, walking, and running. Transitions between diverse arm pos-tures can be inferred from the top left plot of Fig. 2: For $0s \leq t < 5s$, the posture is on idle state. After this interval the user moves her hand up and stabilizes on horizon-tal posture until $t \approx 10s$. For $10s < t < 20s$, the user moves her hand down, stabilizes on idle state and moves her hand up. For $20s \leq t < 27s$, vertical posture is observed, and so on.
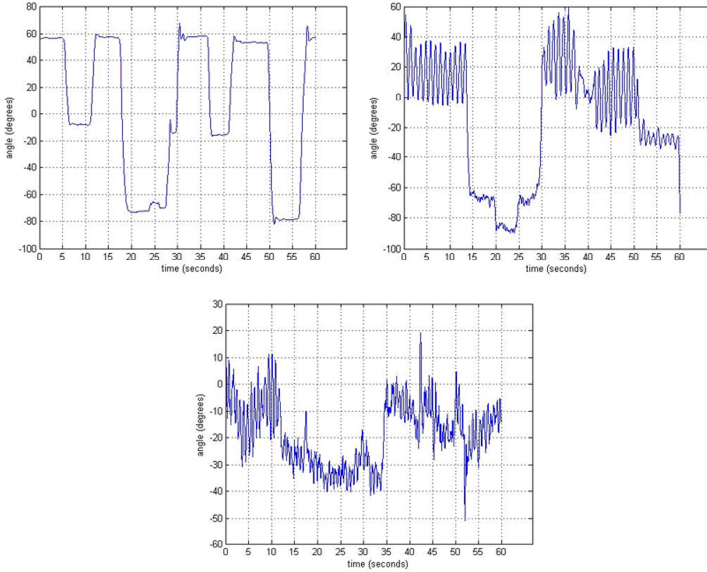
Fig. 2. Pitch Angle Measurements while user is standing (top left), walking (top, right) and running (bottom)

The measurements indicate that with the increase of velocity the noise on the measured signal increases significantly. The noise can be observed on the top right plot of Fig. 2, where the transition from idle posture to horizontal posture is not clearly recognizable at $t \approx 40$. Our current algorithm performs acceptably with users walking with low speed but the accuracy decreases significantly with increased speed due to the high frequency noise introduced into data by walking and running motion.

We implemented a sliding window to detect changes of the hand on pitch angle, α. A window, which contains five angle values obtained in time interval [t-1, t-5], is cre-ated at each time step and upcoming angle is estimated by multiplying them with increasing weights.

$$0.1*\alpha_i + 0.1* \alpha_{i+1} +0.1* \alpha_{i+2} +0.2* \alpha_{i+3} +0.5* \alpha_{i+4} = \alpha_{estimated}. \tag{1}$$

The $\alpha_{estimated}$ angle is compared with the measured angle α_{i+5} to identify if the hand is moving up or down.

$$\alpha_{i+5} > \alpha_{estimated} \rightarrow downside\ change\ . \tag{2}$$

$$\alpha_{i+5} < \alpha_{estimated} \rightarrow upside\ change\ . \tag{3}$$

$$\alpha_{i+5} \approx \alpha_{estimated} \rightarrow no\ change, \begin{cases} \alpha_{i+5} \rightarrow -90°,\ vertical \\ \alpha_{i+5} \rightarrow\ \ \ 0°,\ horizontal\ . \end{cases} \tag{4}$$

However using the pitch angle in one single direction is not sufficient enough to have robust posture recognition. We have also evaluated the case, where the user performs short tilts (rotations around the longitudinal axis) causing an inference on the state transition. For such cases, a filter is implemented on the system which increased the state transition accuracy.

Fig. 3 shows plots of sample pitch and tilt angle measurements of the same motion and corresponding state estimations. For 7s< t <9s, tilt angle is increasing and decreasing instantly (top right plot, Fig. 3) which affects the pitch angle. In spite of the fact that the user holds her hand stable around -20° during angular data measurement, the top left plot of Fig. 3 shows that the pitch angle is changing up to 20°. Same erroneous measurement can be observed for 11s < t < 15s. These unexpected changes cause inaccurate state estimation (bottom left plot, Fig. 3). Therefore, estimation accuracy is increased (bottom right plot, Fig. 3) by introducing the system with a tilt angle filter, which locks the state to the previous one if major changes occur on tilt angles.
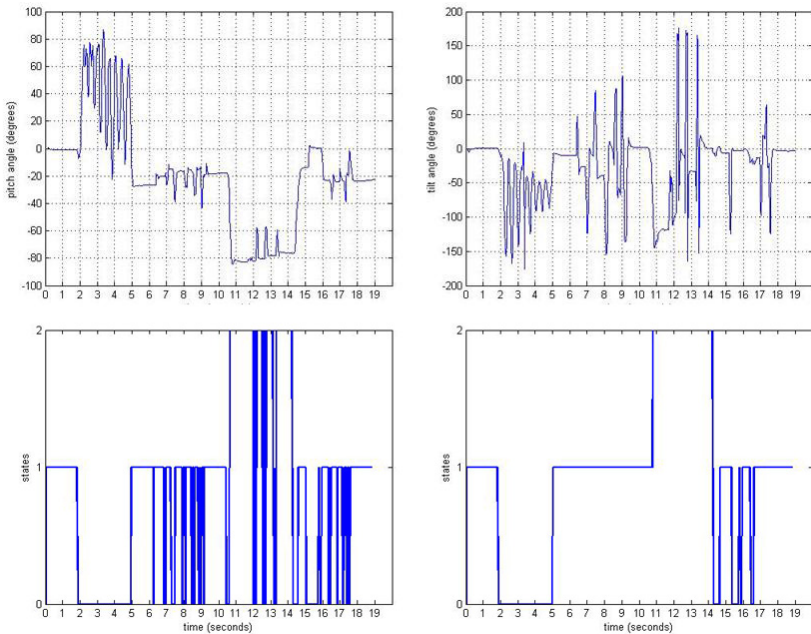
Fig. 3. Sample pitch angle measurement (top left), Tilt measurement of the same posture (top right) causing erroneous estimation – 0: idle state, 1: navigation state, 2: investigation state (bottom left) and increased estimation accuracy with tilt filter (bottom, right)

The system becomes unstable and produces erroneous results when users perform other occasional movement patterns. Therefore we have introduced an additional data, angular velocity, to the recognition system. The change of angular velocity together with the angle allows us more stable recognition results.

Finally we developed a finite state machine to map all possible postures into one of the three states: investigation, navigation and idle (Fig. 4). The investigation state is when a user holds a mobile terminal in vertical position to use it in an Augmented Reality context. In this condition the user needs to investigate point of interest buildings and receives environmental information according to her gaze direction in the local coordinate frame. The navigation state is when a user holds a mobile terminal in horizontal position to use it to render maps or GIS information. Thus, the user receives environmental information in the global coordinate frame. There is a third idle state, where the user is not in either posture and moves her hand freely. In this state, rendering is minimized to save battery power.

The conditions satisfying the state transitions in Fig. 4 are defined in the table in Fig. 5. In this algorithm, the estimated pitch angle value is compared with the angular value perceived from the orientation sensor at that time step. If they are approximately equal, user's arm posture is estimated to be stable and either in investigation or navigation state (3rd and 6th columns of the table in Fig. 5). Other enumerated transitions include conditions which define possible changes between states, i.e. while arm posture is on idle state and the user moves her hand upwards, then it is possible to switch state to navigation or upside change on arm posture continues and state is switched to investigation. The state estimation algorithm is empowered by introducing angular velocity and tilt angle filter to the system.
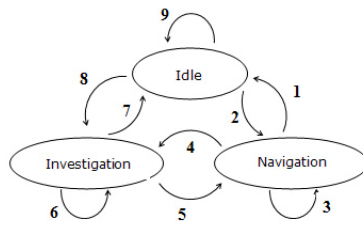
Fig. 4. State transitions between three contexts

1	2	3	4	5	6	7	8	9
$\alpha > 0°$	$\alpha > 0°$	$\alpha \approx 0°$	$\alpha < 0°$	$\alpha < 0°$	$\alpha \approx 90°$	$\alpha > 0°$	$\alpha < 0°$	$\alpha > 0°$
$\omega < -0.75$	$\omega > 0.75$	$\omega \approx 0$	$\omega < -0.75$	$\omega > 0.75$	$\omega \approx 0$	$\omega > 0.75$	$\omega < -0.75$	$|\omega| > 0.75$
$\Delta\alpha < 0$	$\Delta\alpha > 0$	$\Delta\alpha \approx 0$	$\Delta\alpha < 0$	$\Delta\alpha > 0$	$\Delta\alpha \approx 0$	$\Delta\alpha > 0$	$\Delta\alpha < 0$	$\Delta\alpha \neq 0$

$$\omega = angular\ velocity\ (rad/s)$$
$$\Delta\alpha = \alpha_{true} - \alpha_{estimated}$$

Fig. 5. Description of state transitions

Unexpected arm movements of the user can affect the accuracy of the system. While the user holds the PDA in horizontal or vertical position (navigation or investigation state) and suddenly performs fast upward or downward movements with her hand, i.e. waving to somebody, the system is stabilized in the former state with a tolerably accuracy rate.

4   System Setup and Case Study

We have developed an integrated hardware and software system to demonstrate our ideas. We have connected two inertia sensors over RS232 port to a Pocket PC separately. This enabled us to verify our algorithm with more than one sensor technologies. For outdoor case study we also connected a GPS sensor over Bluetooth wireless link (Fig. 6).

The recognition and estimation software is developed in C++. For the case study we have developed an OpenGL|ES based application, also reading the GPS data and orientation data from communication ports. Our OpenGL|ES implementation is based on Vincent Platform [16] and GLUT|ES Windowing libraries [8] (Fig. 7).
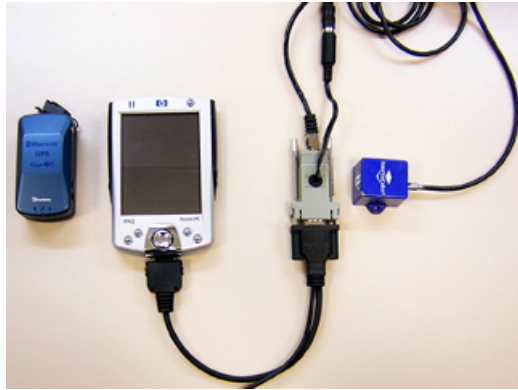
Fig. 6. Hardware prototype: Bluetooth GPS receiver, HP iPAQ Pocket PC h2200 with a RS232 connection to InertiaCube2 (left to right)

We have implemented a prototype navigation system, which guides visitors in downtown Istanbul between Istanbul Biennale venues. However we have experienced serious problems with using both of the orientation sensors we experimented (MTx orientation tracker from Xsens and InterSense's InertiaCube^2) as mobile devices. Both require either large battery packs to operate without a power supply over a long time or a 6V DC power adapter connection. Therefore we have limited the posture recognition experiments in indoors use cases in form of a campus guide (Fig. 8). Note that the investigation mode is not a live video image annotated with data, but a snapshot displayed when the user directs her Pocket PC to the computer science building. For MTX the pitch angle and velocity is acquired from the sensor, for InterSense's InertiaCube^2 we have to compute the angular velocity. For both sensors we experienced stable posture recognition and state switch.
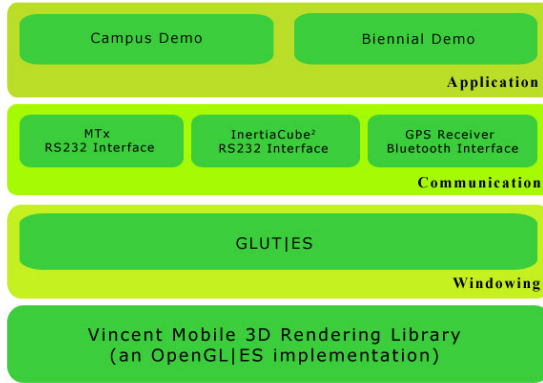
Fig. 7. Software Architecture

We performed a user study to examine the accuracy of our system. In this test, all possible state transitions emphasized in Fig. 4 are performed. The overall accuracy rate is calculated as approximately %87. Performing sudden up-down movements in navigation state and investigation state produced some erroneous results. A usability test covering the prototype's ease of learning, efficiency of use, memorability, error frequency and subjective satisfaction is performed among a group of people aged between 20 to 25. Most participants found the application dependent state transitions appropriate for a pedestrian navigation system, and concluded that once orientation trackers became part of mobile computers they can be used to create natural interaction techniques. Also, the users want to test the prototype outdoors completely with a robust battery pack connection.

Fig. 8. Seamlessly changing interfaces (Demo video can be downloaded from http://graphics.sabanciuniv.edu/Movies/Mobile.avi)

5   Conclusions and Future Work

We have implemented a recognition algorithm with one orientation sensor attached to a Pocket PC to distinguish between two different postures of the hand and an idle state. This data can be used to differentiate between three states to switch between different applications seamlessly. We tested our approach on two different orientation sensors.

In the global coordinate frame, we used GPS sensor data to locate the user, acquire her gaze direction, embed GIS data and provide information about point of interest buildings. In the local coordinate frame, we used orientation sensor data to allow the user interact with the mobile device while performing natural arm postures and perceive information on different user interfaces. By combining these interaction techniques of global and local coordinate frame, we provide a context-aware interaction framework for pedestrian navigation systems on mobile devices by seamlessly changing graphical user interfaces.

Our work shows that once orientation sensors became part of mobile terminals, they can be used to create natural interaction techniques with mobile terminals.

In future we plan to improve our technique to recognize postures while the user is walking and running. Therefore, a robust battery pack connection must be supplied to the system to test the prototype outdoors completely. Also, we want to improve Augmented Reality context of our system by implementing an investigation state with real-time camera input of the environment, where the buildings are annotated with virtual text labels.

References

1. Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a Better Understanding of Context and Context-Awareness. In: Gellersen, H.-W. (ed.) HUC 1999. LNCS, vol. 1707, Springer, Heidelberg (1999)
2. Amft, O., Junker, H., Troster, G.: Detection of eating and drinking arm gestures using intertial body-worn sensors. IEEE Symposium on Wearable Computers (2005)
3. Bao, L., Intille, S.S.: Activity Recognition from User-Annotated Acceleration Data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
4. Burigat, S., Chittaro, L.: Location-aware visualization of VRML models in GPS-based mobile guides. In: Web3D '05: Proceedings of the tenth international conference on 3D Web technology, pp. 57–64. ACM Press, New York (2005)
5. Cheng, Y., Chawathe, Y., Lamarca, A., Krumm, J.: Accuracy characterization for metropolitanscale wi-fi localization. Third International Conference on Mobile Systems, Applications and Services (2005)
6. Feiner, S., MacIntyre, B., Höllerer, T., Webster, A.: A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. In: Proc. ISWC'97. Cambridge, MA, USA, pp. 74–81 (1997)
7. Foxlin, E.: Pedestrian Tracking with Shoe-Mounted Intertial Sensors. IEEE Computer Graphics and Applications, Nov/Dec (2005)
8. GLUT|ES - The OpenGL|ES Utility Toolkit, (as of April 2006) http://glutes.sourceforge.net/

9. Howard, A., Siddiqi, S., Sukhatme, G.S.: An experimental study of localization using wireless ethernet. In: Proceedings of the 4th International Conference on Field and Service Robotics, Lake Yamanaka, Japan (2003)
10. Lee, S., Mase, K.: Activity and Location Recognition Using Wearable Sensors. IEEE Pervasive Computing 1(3), 24–32 (2002)
11. Lee, S., Mase, K.: Recognition of Walking Behaviors for Pedestrian Navigation. In: Proc. of IEEE International Conference on Control Applications, Mexico City, Mexico (2001)
12. Peternier, A., Vexo, F., Thalmann, D.: Wearable Mixed Reality System In Less Than 1 Pound. Eurographics Symposium on Virtual Environments (2006)
13. Rohs, M.: Real-world interaction with camera-phones. In: International Symposium on Ubiquitous Computing Systems (2004)
14. Schmalstieg, D., Reitmayr, G.: Augmented Reality as a Medium for Cartography. In: Cartwright, W., Peterson, M., Gartner, G. (eds.) Multimedia Cartography, 2nd edn. Springer, Heidelberg (2006)
15. Shimada, S., Tanizaki, M., Maruyama, K.: Ubiquitous Spatial-Information Services Using Cell Phones. IEEE Micro (2002)
16. Vincent Mobile 3D Rendering Library (as of April 2006),
    http://sourceforge.net/projects/ogl-es
17. Wagner, D., Schmalstieg, D.: First Steps Towards Handheld Augmented Reality. In: Proc. Seventh IEEE International Symposium on Wearable Computers (2003)
18. Widgor, D., Balakrishnan, R.: TiltText: Using Tilt for Text Input to Mobile Phones, UIST 2003, CHI Letters, vol. 5(2) (2003)