An Exploration of Mobile Augmented Reality

Zach Gildersleeve

University of Utah – CS 6360 Virtual Reality

ABSTRACT

Early implementations of augmented reality (AR) have been limited by tracking and display technology – the same encumbrances that often restrict virtual reality (VR) research. Likewise, AR applications are often confined indoors. Approaching AR from a mobile computing framework using consumer products can be a cheaper, more practical solution to applying AR to outdoor applications, but the use of consumer products brings a new set of limitations. This paper presents an AR technique using a handheld GPS unit, consumer laptop, and prosumer digital video camera, and illustrates some possible applications of mobile AR technology.

Keywords: Augmented Reality, GPS Navigation, Mobile Computing, Gaming, Virtual Environments

1 INTRODUCTION

With the rise of affordable mobile computing technology such as handheld GPS units and camera equipped cell phones, the possibility of augmented reality applications at the consumer level is increasing. GPS allows straightforward geospatial positioning, but ultimately is limited by an error range that falls within the range of interest for typical AR applications. The use of a hybrid tracking system is usually suggested in mobile augmented reality systems to control this error, but such hybrid tracking systems are sources of additional error, and are limited by the resolution of displays and the measurement of devices. At the global scale, a small difference in measurement of an angle or position can translate to an error of hundreds of meters at sufficient distance, distances that contain points of interest relevant to the application of AR.

This paper describes a system to interface GPS positional data within an AR system, and examines the benefits and problems associated with such a system. In doing so, the system will demonstrate a large scale navigation utility, for identifying points of interests on a city scale, including mountain peaks, civil buildings, and airports. Such a system in theory could supplement a critical step in GIS and related utilities. Image data can be geospatially located in terms of where an image was originally captured, but currently the image itself provides no data as to the specific orientation of the camera, and what points of interests are visible in the images.

2 RELATED WORK

Previous research in mobile (outdoor) augmented reality has typically focused on navigation of large spaces, or in gaming applications. Gaming in particular tends to push the frontier of AR, and is essentially a specific application of navigation. Feiner et al designed a campus navigation system using a see-through head mounted display (HMD) with built-in orientation tracker, differential GPS, and handheld computer [5]. Using consumer products available in the late 1990's, their system weighed 40 pounds. Azuma [2,3] discusses tracking technology available in mobile outdoor augmented reality, including GPS, inertial, and passive optical systems borrowing from computer vision techniques, and suggests that any AR application would require more than one form of tracking.

An augmented reality outdoor gaming system was introduced in ARQuake [1], which is built on ARToolKit [8] using GPS tracking. ARQuake uses non-visible architectural models tracked and aligned with real world building to provide occlusion of ARQuake monsters, thus exact registration is important to seamlessly integrating the virtual world. Thomas detailed that GPS alone was sufficient for distance of greater than 50m, and suggests that ideal registration is a function of scale and distance [9].

Cheok, with Human Pacman, introduced social gaming to mobile AR by running a virtual environment game server and wirelessly feeding the actions of other users into an AR HMD [4]. An examination of further related work suggests that as wireless networks and data storage capabilities grow, the computing hardware required to implement mobile AR systems reduces down to only a tracker and display, as seen in GPS::Tron [6].

3 SYSTEM DESIGN

The AR technique implemented and discussed below uses some standard consumer level technology:

- Apple Powerbook Laptop
- Garmin eTrex Legend CX handheld GPS
- Panasonic DVX-100 firewire camera

The one missing – but necessary – component is an orientation tracker, and thus no data is available regarding the rotation of the camera in 3D space. Therefore, this technique assumes the camera is level with the horizon (provided by a bubble levelled tripod). Panning rotation around the camera's *y*-axis is ignored and substituted for either the bearing of travel when moving, or user input when stationary. Optionally, such as in ARQuake [1], this information could be provided via ARToolKit markers.

3.1 ARToolKit Implementation

To keep this system open ended within AR, it is built using ARToolKit video capture, markers, and OpenGL routines. To incorporate GPS locations within ARToolKit, a shared coordinate system is necessary. The implemented hybrid approach uses GPS provided latitude, longitude, and altitude positions as surrogate markers, but forces ARToolKit to treat these pseudomarkers as real markers by injecting them into the ARToolKit pipeline.

In the standard pipeline, ARToolKit uses image processing techniques to locate markers within each frame, determines the corner vertices and orientation of the marker, and passes this information to a function that computes an inverse projection matrix for each marker, thus arriving at the camera's position relative to the markers.

With this system, the first step is precomputed, each pseudomarker is assumed to be orientated aligned and facing the camera. As detailed in Section 4, the size of the pseudomarker is determined as a function of the distance to the camera, and by calculating in ARToolKit frame coordinates where each pseudomarker would be if it actually existed, a projection matrix can be accurately created.

3.2 GPS Interface

Access to the current GPS provided position, expressed in latitude, longitude, and altitude is provided via GPSBabel [7]. GPSBabel is typically used for the transfer of waypoints, routes, and tracks, standard offline GPS information, although recent extensions allow for real time positional data. The depth of GPSBabel testifies to the proprietary and diverse nature of the consumer GPS market, which is one roadblock to easy assimilation of AR technology. GPS is provided free to civilians as a good will service; most open source GPS interfaces are the product of reverse engineering classified systems.

In the implemented system, current latitude and longitude can be returned roughly every 0.1 seconds, while full 3D position including altitude takes on the average 1 second. This time disparity is due to the used GPS make and model, other devices would experience similar, but distinct poll times.

3.3 Video Camera

The digital video camera used is admittedly overkill, but sidesteps several encountered problems regarding use of the GPS unit and smaller webcam unit at the same time (USB conflict) and offers more control during use in the varied lighting conditions met outside. The use of this camera did limit mobility.

4 Coordinate Registration

To force ARToolKit to treat the GPS pseudomarkers as real markers, the pseudomarker must be located in the ARToolKit captured image frame coordinates. To do this, we need both the distance and orientation from the camera (and GPS unit) to the pseudomarker. This information will be encoded into ARToolKit's ARMarkerInfo data structure.

4.1 Pseudomarker Frame Coordinate

The distance between two geospatial locations is not a trivial problem due to the unevenness of the earth's radius, and the loss of decimal precision experienced at small distances. We took advantage of the estimation offered by the transformation of latitude A and longitude B, where R is 6371km:

$$[x , y , z] = [R\cos(A)\cos(B) , R\cos(A) \sin(B) , R\sin(A)] \quad (1)$$

to use the familiar distance equation:

$$d = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \quad (2)$$

The bearing between two points [lat1, lon1] and [lat2, lon2], can be expressed in radians by the following equation:

$$bear = \operatorname{atan2}(\sin(lon2 - lon1)\cos(lat2), \cos(lat1)\sin(lat2) - \sin(lat1)\cos(lat2)\cos(lon2 - lon1)) \,\%\, (2\Pi) \quad (3)$$

These two sets of equations form the basis for locating a pseudomarker in frame coordinates. From this point, a series of similar triangles narrows the field of view from a global angle to the camera's field of view and finally to the coordinate x offset from the center of the frame. This process is repeated for the pseudomarker's frame y offset using the difference in altitude.

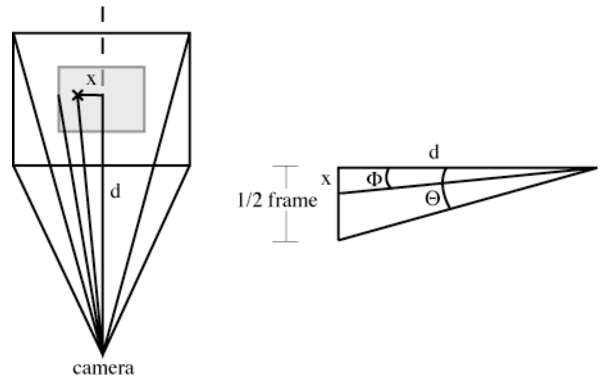

Figure 1. Similar triangles identify pseudomarker frame coordinate

In Figure 1, we find the unknown x offset via the known frame dimension, the camera's field of view angle Θ, and Φ is the bearing angle towards the pseudomarker based on the defined rotation as mentioned in Section 3.

4.2 Pseudomarker Scaling

By manipulating the GPS located pseudomarker into frame coordinates, the ARToolKit marker structure can be built appropriately. As mentioned, the scale of the pseudomarker is a function of distance d. Each pseudomarker is assumed to appear 10m x 10m at a distance from the camera of 10m. Pseudomarkers at distances under 1km are scaled naturally via the inverse square law, and this scaling is directly built into the markers. However, pseudomarkers at distances farther than 1km from the camera are scaled at a slower linear rate, and this scaling is done in the OpenGL draw routine. This is to avoid AR objects associated with pseudomarkers collapsing into subpixel area too quickly, and thus to augment the view from the camera beyond what the camera's resolution might be limited to. In practice, though, any AR system is limited by its display capabilities, and this system is no exception.

5 Testing Situation and Results

The implemented system was limited by three factors:

- No orientation data
- Cumbersome setup
- Error range of GPS data

The first two limitations were overcome by deciding to limit the camera setup to a static tripod. This way, the camera could be orientated using a standard analogue compass around the camera's y-axis, and the AR world's orientation could match the cameras. This simulates the data provided via an orientation tracking device.

Multiple geospatial points of interest (mountain peaks, civil buildings, hospitals, freeway interchanges) were incorporated as pseudomarkers. Each pseudomarker has an associated position, scale, and text label, and the exact point is represented by a glutSolidCone. For static work, the GPS could be polled only once for the initial current position, and by reusing this data dramatically increase frame rate.

As mentioned before, the camera was aligned using an analogue compass, and this information was used as a starting bearing for both the camera and the pseudomarker bearing calculations.

zgildersleeve@gmail.com

Figure 2. Results from camera location *A* with a bearing of 240°

Figure 2 illustrates typical results, with the vertical dashed lines added afterwards to help visualize where the actual points of interest are in the frame. In Figure 2, points of interest around 2-20km from the camera show accurate registration, while points on the mountain range to the left experience error as a results of error compounded over distance, and limited precision of measuring necessary camera angles.

Figure 3. Results from camera location *B* with bearing of 40°

Figure 4. Results from camera location *A* with bearing of 0°

Figures 3 and 4 together illustrates identifying the same point of interest from two different camera locations and orientations. These points are 2-5km from the camera, depending on camera position *A* or *B*.

Figure 5. Incremented latitude markers

Figure 5 displays a part of 100 markers set at 1.0e-5 increments in latitude. Towards the left, as markers exceed roughly .050km in distance from the camera, the registration experiences less error and the markers are aligned and distributed correctly. The goal is roughly 15m from the camera, inside what was found to be the error margins for the implemented technique, and the markers experience severe error here compared to the uniform distribution at farther distances. Any points of interest within this range would need to be supplemented with a hybrid tracking system for exact registration.

6 Conclusions and Future Work

While the resulting images are admittedly somewhat difficult to decipher, and lack the overall robustness that an orientation tracker would bring, they do illustrate the successes and pitfalls of mobile augmented reality systems. GPS provides location data on the global, mostly 2D, scale, which is adequate for 2D applications such as GPS::Tron, but is somewhat limited in full 3D. This system was most accurate when the point of interest being tracked was in the range 20m – 10km; anything under this range falls within the error and noise of GPS positioning and loss of numerical precision, while larger distances tend to over-rely on accurate angle measurements that can introduce significant error at large distances.

Using ARToolKit is one solution to using the hybrid tracking technique suggested in prior research, but ultimately is probably not the best option. ARToolKit and the implemented system are commonly limited by reliance in camera resolution and calibration, as well as accurate measurements when aligning markers and pseudomarkers with the real world. Current research indicates movement towards using computer vision techniques, wireless networking, and more available differential GPS systems to fine tune the positional data provided by GPS, and to supplement the orientation information offered by a suitable tracker. The use of a high-resolution digital still camera, either with a similar real-time system or an offline GIS style system also is promising for overcoming the resolution barrier, and for the external data information available with modern images.

Reference

[1] Avery, B., Thomas, B. H., Velikovsky, J., and Piekarski, W. 2005. Outdoor augmented reality gaming on five dollars a day. In *Proceedings of the Sixth Australasian Conference on User interface*

- *Volume 40* (Newcastle, Australia, January 30 - February 03, 2005). M. Billinghurst and A. Cockburn, Eds. ACM International Conference Proceeding Series, vol. 104. Australian Computer Society, Darlinghurst, Australia, 79-88.

[2] Azuma, R. T., et al. Making Augmented Reality Work Outdoors Requires Hybrid Tracking. In 1st Int'l Workshop on Augmented Reality, San Francisco, Ca, Nov 1998.

[3] Azuma, R. T. (1999): The Challenge of Making Augmented Reality Work Outdoors. In Mixed Reality: Merging Real and Virtual Worlds, pp 379-390, Mar 1999.

[4] Cheok, A., Goh, K. H., Liu, W., Farbiz, F., Fong, S. W., Teo, S. L., Li, Y., and Yang, X. 2004. Human Pacman: a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. *Personal Ubiquitous Comput*. 8, 2 (May. 2004), 71-81.

[5] Feiner, S., MacIntyre, B., Hollerer, T., and Webster, A. 1997. A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. In *Proceedings of the 1st IEEE international Symposium on Wearable Computers* (October 13 - 14, 1997). ISWC. IEEE Computer Society, Washington, DC, 74.

[6] GPS::Tron http://gps-tron.datenmafia.org/

[7] GPSBabel http://www.gpsbabel.org/

[8] Kato, H., Billinghurst, M. (1999) Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System. In Proceedings of the 2nd International Workshop on Augmented Reality (IWAR 99). October, San Francisco, USA.

[9] Thomas, B., Close, B., Donoghue, J., Squires, J., Bondi, P., and Piekarski, W. 2002. First Person Indoor/Outdoor Augmented Reality Application: ARQuake. *Personal Ubiquitous Comput*. 6, 1 (Jan. 2002), 75-86.