Rapid and Accurate Deployment of Fiducial Markers for Augmented Reality

Gerhard Schall, Joseph Newman and Dieter Schmalstieg
Institute for Computer Graphics and Vision, Graz University of Technology
Inffeldgasse 16a, A-8010 Graz, Austria
phone : +43 316 873-5011, fax : +43 316 873-5050
`e-mail: {schall, newman, schmalstieg}@icg.tu-graz.ac.at`

Abstract

Fiducial markers are commonly used for providing the tracking required by Augmented Reality applications. Libraries exist that accurately track markers using cheap cameras, whilst the cost of paper and toner associated with the manufacture of the markers is negligible. Attempts have been made at deploying markers over a wide area in order to extend tracking range. This paper describes techniques for rapidly and accurately surveying the locations of widely distributed markers, whilst simultaneously building a model of the environment.

1 Introduction and Motivation

Augmented Reality (AR) has the potential to provide a natural interface to the “calm” pervasive technology anticipated in large-scale Ubiquitous Computing [20] environments. However, most AR applications have hitherto been constrained, by the working volumes of tracking technologies, to static spaces of a few cubic metres, such as the Boeing wire assembly [5] example. Systems aiming at mobility like the Touring machine [7], Sentient AR [15], or Tinmith [17] have relied on wide-area trackers, such as GPS, which provide modest levels of accuracy at low update rates. Sentient Computing [2, 8, 9] demands a detailed and up-to-date model of spatial relationships, which appear to reproduce the perceptions a user has of the world. Given a suitably large model it is possible to create AR scenarios that allow users to roam through a wide area interacting with a rich and responsive environment. Hence the two major ingredients necessary to experiment with and build “Location-aware” AR applications, are tracking and modelling.

Many research groups find the expense of purchasing and installing high performance tracking equipment to be prohibitive. Consequently, fiducial tracking libraries such as ARToolkit [11] have become very popular, allowing anyone with a PC, webcam and printer to become Mixed and Augmented Reality researchers or developers. The library was primarily intended to be used in a video-see-through mode, registering superimposed 3D graphics on top of the fiducial markers visible in video streams. However, a number of groups have taken a different tack, using the library as a general-purpose tracking algorithm [10].

The decision to use fiducial markers for the purposes of tracking is an entirely pragmatic one. Clearly, a natural feature vision-based tracker [3, 21] would be more desirable as it would not require visual pollution of the environment. Also, advanced vision algorithms are increasingly capable of performing Simultaneous Localisation and Mapping (SLAM) [6] effectively building a model on the fly. However, even the most succesful — and often proprietary [1] — Computer Vision research solutions have yet to prove themselves sufficiently robust as drivers of AR experiences in real environments. Nevertheless, by performing a high quality survey of fiducial markers and their environs using the most sophisticated geodetic equipment, we can establish a ground truth against which vision techniques can be evaluated. Yet another line of experimentation would anticipate future environments distributed sparsely with fiducials being used to initialise, and stabilise vision trackers.

2 Spatial Relationships

Spatial relationships can, in general, be represented by a graph [4], in which objects are nodes, and spatial relationships between objects are directed edges. Each edge represents the spatial transformation between objects. A complete spatial relationship (SR) graph would represent environmental state in its entirety and could be used to query relationships between two arbitrary objects. This approach was developed within the course of the Ubiquitous Tracking (Ubitrack) [16] project, and was intended to model dynamic relationships. The survey is (or should be) a static problem, but still lends itself to the same approach.

We can only make *estimates* of geometric relationships between real objects by making *measurements*. In the case of the survey, each measurement yields a 3 DOF geometric relationship describing a point in 3D space but corrupted by noise. The quality of a measurement is described using a set of *attributes* would ordinarily include properties such as latency or a confidence value, but in this case would best be represented by a 3×3 covariance matrix. A *tracker* is a *sensor* that takes measurements of the spatial relationship between itself and other objects or *locatables*. Thus, edges are added to the graph or attributes updated in existing edges. In our case, the Leica TPS 700 Total Station [13] acts as a kind of tracker, along with head-mounted cameras. A fiducial marker acts as a locatable.

Example: There are three objects of interest in the optical shared tracking scenario [12] shown in figure 1, namely a single fixed marker and two cameras. A graph depicting this situation would be complete and all edges have attributes determined solely by known spatial relationships. If the leftmost person wishes to augment their view of the rightmost, then a query for the spacial relationship q^{app}_{DA} is issued. The solution involves the concatentation of two other spatial relationships: q^{e}_{AB} and the inverse of q^{e}_{DB}. As the graph is directed the inversion operation results in a new edge q^{inv}_{BD}. The absolute pose of marker B, and the confidence with which it is known, is irrelevant; as long as the marker is visible to both parties. However, a dynamic map application querying the absolute poses of cameras A and D relative to the world origin, clearly requires more knowledge about marker B. This issue will be revisited in section 5.

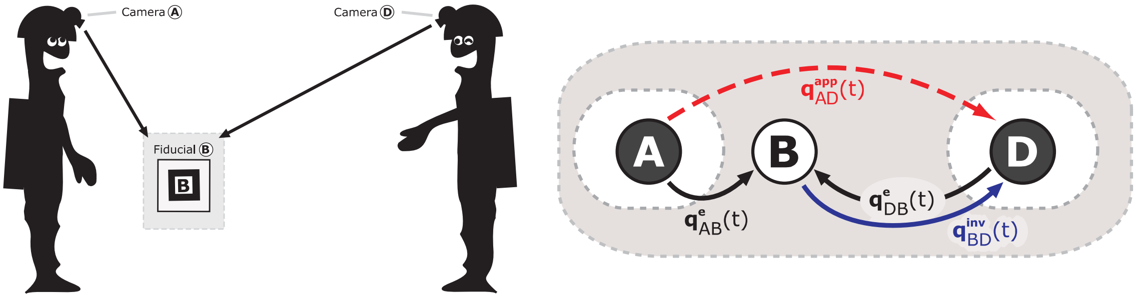

Figure 1. Example setup: Both cameras A and D detect fixed fiducial marker B. The application is interested in the pose of camera D relative to the pose of camera A (q^{app}_{DA}), which can be calculated by inverting the relationship $q^{e}_{DB}(= q^{inv}_{BD})$ and concatenating with q^{e}_{AB}.

3 Workflow

In general the surveying process workflow summarised in Figure 2 consists of the following steps:

1. Firstly, preparations for surveying are undertaken. These mainly consist of making a preliminary examination of the surveying area and attaching marker templates on the walls at a maximum interval of two metres.

2. The surveying area is then divided into measurable segments. That is the areas which can be measured from wherever the Total station is positioned. At least one measurement is necessary for each room if there is a clear line of sight to the important surfaces and edges. Note that it is important that neighbouring surveying areas overlap, ensuring that common points of correspondence (so called “pairings”), occur in both adjacent areas, acting as a bridge. A minimum of three correspondences with good geometry (i.e. neither co-linear nor too close together) are necessary in order to be able to calculate the transformation from the coordinate reference frame in one room and the coordinate reference frame in the other room. For each marker template the four corners of the square patch of the marker are measured, and the position and orientation calculated.

3. Steps 3 to 5 are performed iteratively for all parts of the surveying area. All points of the room geometry are measured using the Total station, followed by the edges of portals and marker templates and at least the pairings to the neighbouring room or corridor. Ids are given to all measured points, thus ensuring that each point can be uniquely identified and referencedwhen building the 3D model.

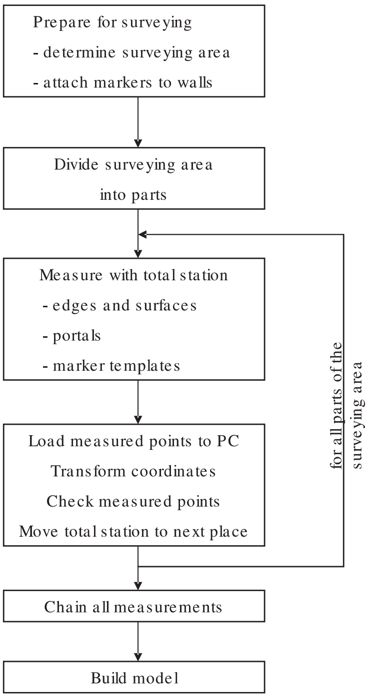

Figure 2. Surveying and modelling workflow

4. The acquired data is transferred from the Total Station to the PC. To assure the measurements have been conducted successfully, all points are checked.

5. The measured points are transformed from polar into cartesian coordinate system. Scripts parse the original measurement file and converts the points into cartesian coordinates and saves them in a Matlab file format, which is for instance illustrated graphically in Figure 6.

6. A script transforms the points from the Matlab file format into BAUML [18] (Building Augmentation Markup Language) file format. BAUML is an XML language for the representation of geometric information. It allows the building geometry (e.g. walls, floors and corridors) to be stored, as well as the positions of markers. Due to the recursive definition of the language a tree structure of spatial objects, where objects are composed of a number of smaller objects can be created.

7. Having completed these steps the Total Station is moved to the next measurement position and steps 3 to 5 are now repeated in the next part of the surveying area until the entire area is surveyed.

8. Finally, all the measured points are transformed into one common reference frame. They are then chained together and merged, yielding a 3D model of thesurveying area.

Figure 4 depicts a simplified sketch of how measurements of neighbouring surveying areas are conducted. At position A all points of measurement area A including the three pairing points

$P1$, $P2$ and $P3$ are measured. The Total Station is then moved to position B, where the points in area B are similarly determined. Note that the points in both areas are measured in different coordinate system.

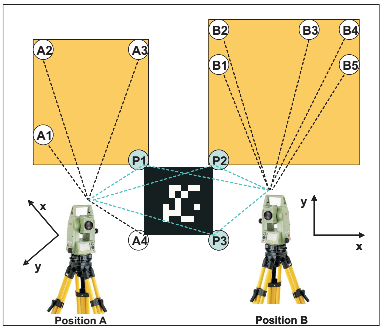

Figure 3. Bridging of adjacent surveying areas

Figure 4 shows a graph representation of the measurements illustrated in Figure 3. The magenta circles represent the coordinate systems of the Total Station where they are placed in positions A and B. The pairings, consisting of common points in the overlapping surveying area, are denoted by cyan nodes. In this case the high contrast corners of the fiducial marker are used to provide the minimal set of three, appropriately conditioned, common points required to calculate the transformation indicated by the magenta arrow from A to B. The remaining white circles represent the other measured points on edges and surfaces that are to be included in the model. It should then be possible to transform all the points measured from location A into the coordinate system of location B. All the points are now in the same reference frame and can be chained to produce an comprehensive model.

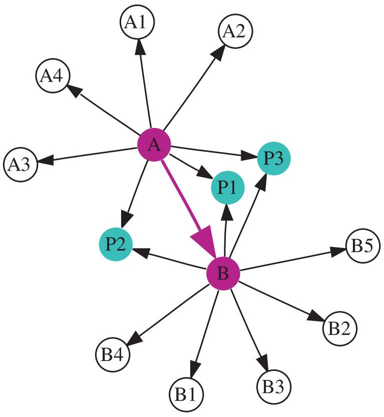

Figure 4. Spatial relationships involved in the bridging of adjacent surveying areas

Figure 5 shows a subgraph obtained from the survey performed at our institute. The magenta coloured circles represent the measurement positions A to D of the total station. The pairings of each part of the surveying area are printed in cyan. The white nodes represent points measured on edges and surfaces. Using the appropriate pairings all the points measured at location A are transformed into the coordinate system of location B of the total station. Next, all points from location A and B are transformed to the reference frame of location C and consequently all parts of the surveying area are transformed into the same coordinate system of location D.

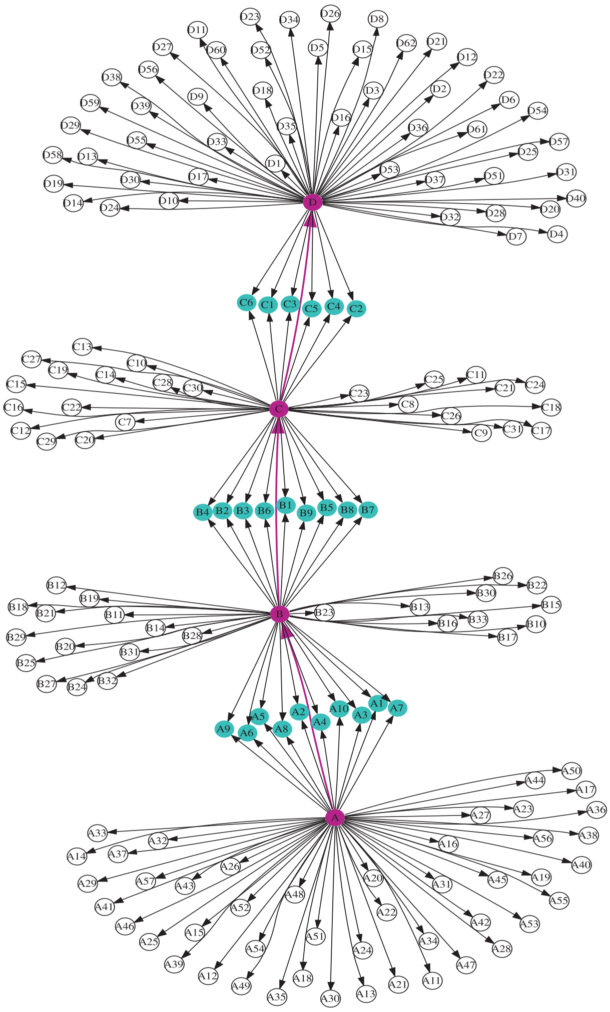

Figure 5. Graph of inter-relationships between points; their respective coordinate frames; and the transformations between coordinate frames

4 Results

The quantitative results of the survey can be seen in Figure 6 where the vertices, drawn by Matlab, already form a discernible structure. The unhighlighted black crosses in this figure correspond to the white circles in Figure 5,whilst the crosses highlighted in cyan are the pairings indicated by cyan circles in Figure 5. Magenta coloured circles in both diagrams represent the locations A through D where the Total Station was placed. Figure 7 shows the qualitative effectiveness of the approach, in which the floor, walls, portals (doors) and ceiling are clearly visible. The positions and orientations of the fiducial markers are represented as red cones.

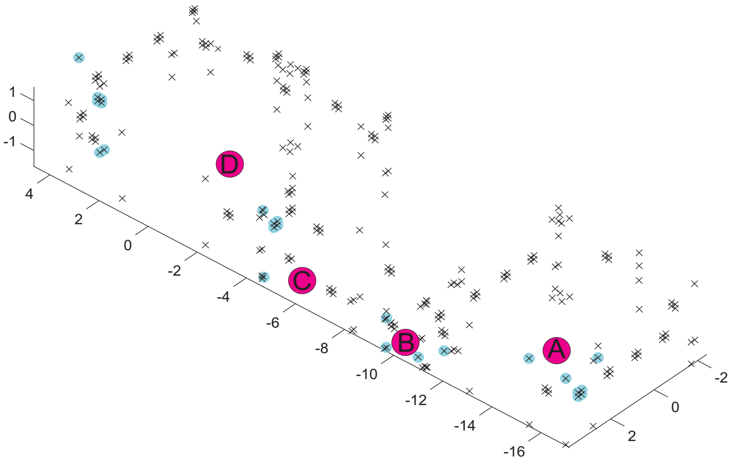

Figure 6. Cloud of surveyed points in which structure of walls can be discerned.

5 Future Work

Consideration of Error Figure 8 illustrates a similar situation to that described in the example in section 2. Camera A detects marker B, whilst camera D detects marker C. Markers B and C are fixed relative to the world origin X. When the leftmost person wishes to augment their partner, a more complex situation emerges. In order to calculate the spatial relationship q_{DB}^{app} the relations q_{BX}^{e} and q_{DC}^{e} must be inverted yielding new relationships q_{BX}^{inv} and q_{CD}^{inv}. The subsequent concatenation can be expressed:

$$q_{DB}^{app} = q_{AB}^{e} \rightarrow q_{XB}^{inv} \rightarrow q_{XC}^{e} \rightarrow q_{CD}^{inv}$$

Inversion The inversion process depends greatly on how uncertainty is measured. Following the lead of Drummond et al. we intend to use an exponential map to represent all 6 DOF transformations. As described earlier in section 2 all relationships have attributes, which in this case we can model as a 6×6 covariance matrix.

Given a transformation represented by a matrix E which has an associated covariance matrix C, the covariance matrix C' associated with the inverse of E is given by:

$$C = \mathrm{Ad}(E) C' \mathrm{Ad}(E)^T$$

Where Ad is the action of E over the Lie algebra by conjugation.

Propogation Similarly, given two transformations with covariances C_1 and C_2 the covariance associated with an optimal concatenation is given by:

$$C_{opt} = C_1 (C_1 + C_2)^{-1} C_2$$

For further details see Drummond on representing, propagating and combining PDFs [19].

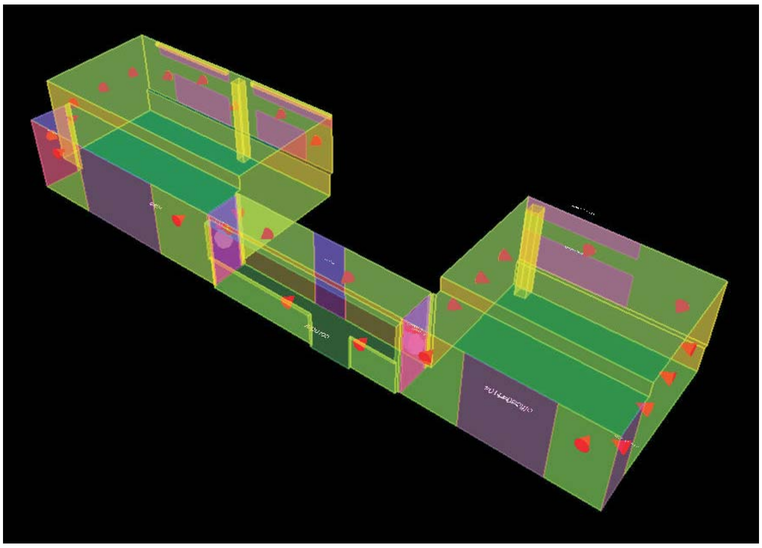

Figure 7. Screenshot of BAUML viewer visualising a section of the building model

Fiducial pose uncertainty It is not yet clear how to represent the uncertainty from ARToolkit. A confidence value is associated with the probability that the marker is the actual one that has been detected and recent work has characterised the error of the reconstructed pose [14]. Usable heuristics must be developed to assign uncertainties to each pose estimate.

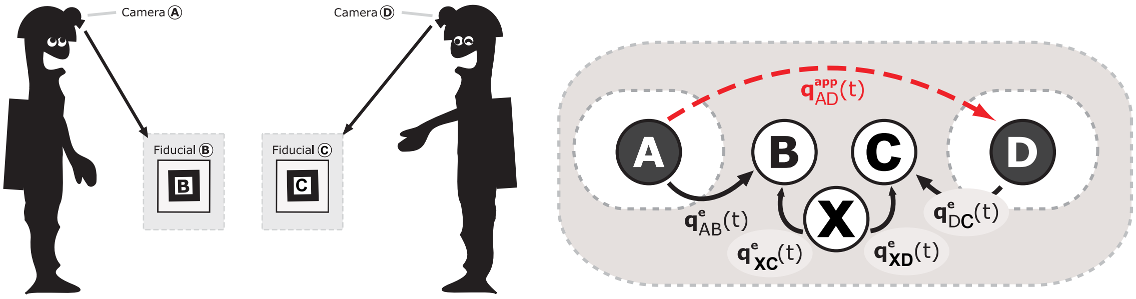

Figure 8. Example setup: Cameras A and D detect fiducial markers B and C respectively. The fiducial markers are surveyed and are related to the "root" world node, X.

Bundle adjustment Given that the location of many of the points may be measured multiple times, it becomes possible to apply a technique called *bundle adjustment* to obtain optimal (non-linear) least-squares estimates for the positions of each point, and the poses of each fiducial marker. This process should take into account the different accuracies associated with each measurement, as the range accuracy of the Total Station decreases with increasing distance.

6 Conclusions

An approach to surveying the locations of fiducial markers has been demonstrated. A general model of the environment is also constructed simultaneously which is sufficiently accurate to act as a ground truth against which to compare SLAM experiments. It is also possiblet to apply the location information to the inference of context in Ubiquitous Computing applications. Some of the subtleties involved in handling the resultant data has been discussed; techniques demonstrated; and further solutions proposed.

References

[1] 2d3. Boujou bullet. http://www.2d3.com/jsp/products/product-overview.jsp?pro duct=7 [2004, December 20], 2004. 2

[2] Mike Addlesee, Rupert Curwen, Steve Hodges, Joseph Newman, Pete Steggles, Andy Ward, and Andy Hopper. Implementing a sentient computing system. *IEEE Computer*, 34(8):50–56, 2001. 1

[3] R. Azuma and G. Bishop. Improving static and dynamic registration in an optical see-through HMD. In *Proc. Siggraph'94*, pages 194 – 204, Orlando, July 1994. 2

[4] B. Brumitt, J. Krumm, B. Meyers, and S. Shafer. Ubiquitous computing and the role of geometry. *IEEE Personal Communications*, pages 41–43, October 2000. 2

[5] D. Curtis, D. Mizell, P. Gruenbaum, and A. Janin. Several devils in the details: Making an AR app work in the airplane factory. In *First IEEE Workshop on Augmented Reality*, November 1998. 1

[6] A.J. Davison, W.W. Mayol, and D.W. Murray. Real-time localisation and mapping with wearable active visi on. In *Proc. IEEE International Symposium on Mixed and Augmented Reality*. IEEE Computer Society Press, October 2003. (To appear). 2

[7] S. Feiner, B. MacIntyre, T. Höllerer, and T. Webster. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. In *Proc. of ISWC 1997*, October 1997. 1

[8] Andy Harter, Andy Hopper, Pete Steggles, Andy Ward, and Paul Webster. The anatomy of a context-aware application. In *Mobile Computing and Networking*, pages 59–68, 1999. 1

[9] Jeffrey Hightower, Barry Brumitt, and Gaetano Borriello. The location stack: A layered model for location in ubiquitous computing. In *Proceedings of the 4th IEEE Workshop on Mobile Computing Systems & Applications (WMCSA 2002)*, pages 22–28, Callicoon, NY, June 2002. 1

[10] Michael Kalkusch, Thomas Lidy, Michael Knapp, Gerhard Reitmayr, Hannes Kaufmann, and Dieter Schmalstieg. Structured visual markers for indoor pathfinding. In *Proceedings of the First IEEE International Workshop on ARToolKit (ART02)*, Darmstadt, October 2002. IEEE Computer Society. 2

[11] H. Kato and Mark Billinghurst. Marker tracking and HMD calibration for a video-based augmented reality conferenencing system. In *Proc. IWAR'99*, pages 85–94, San Francisco, CA, USA, October 21–22 1999. IEEE CS. 1

[12] Florian Ledermann, Gerhard Reitmayr, and Dieter Schmalstieg. Dynamically shared optical tracking. In *Proceedings of the First Augmented Reality Toolkit Workshop*, 2002. 2

[13] Leica Geosystems. Advanced realtime tracking gmbh. http://www.leica-geosystems.com/corporate/en/products/total_stations/lgs_5166.htm [2004, December 20], 2004. 2

[14] P. Malbezin, W. Piekarski, and B.H. Thomas. Measuring artoolkit accuracy in long distance tracking experiments. In *1st Int'l Augmented Reality Toolkit Workshop*, Darmstadt, Germany, 2002. 8

[15] Joseph Newman, David Ingram, and Andy Hopper. Augmented reality in a wide area sentient environment. In *Proc. of IEEE and ACM Int. Symp. on Augmented Reality (ISAR 2001)*, pages 77–86, New York, NY, October 2001. 1

[16] Joseph Newman, Martin Wagner, Martin Bauer, Asa MacWi lliams, Thomas Pintaric, Dagmar Beyer, Daniel Pustka, Franz Strasser, Dieter Schmalstieg, and Gudrun Klinker. Ubiquitous tracking for augmented reality. In *Proceedings of ISMAR 2004*, November 2004. 2

[17] Wayne Piekarski and Bruce H. Thomas. Tinmith-evo5: A software architecture for supporting research into outdoor augmented reality environments. Technical report, Wearable Computer Laboratory, University of South Australia, December 2001. 1

[18] Gerhard Reitmayr. A software design for augmented reality. Dissertation, 2004. 4

[19] Paul Smith, Tom Drummond, and Kimon Rousoppoulos. Computing map trajectories by representing, propagating and combining pdfs over groups. pages 1275–1282, Nice, France, 2003. 7

[20] Mark Weiser and John Seely Brown. *Beyond Calculation: The Next Fifty Years of Computing*. Copernicus, 1998. 1

[21] Suya You, Ulrich Neumann, and Ronald Azuma. Hybrid inertial and vision tracking for augmented reality registration. In *Proceedings of the IEEE Virtual Reality*, page 260. IEEE Computer Society, 1999. 2