Better Living Through Geometry

Barry Brumitt, Steven Shafer
{barry, stevensh}@microsoft.com
Microsoft Research

Mark Weiser described ubiquitous computing as, "invisible, everywhere computing that does not live on a personal device of any sort, but is in the woodwork everywhere."[8] The EasyLiving project at Microsoft Research is focused on those aspects of ubiquitous computing relevant to smart environments, including work in distributed computing, geometric world modeling, computer vision, and user interfaces. Though the need for research in distributed computing, perception, and interfaces is widely recognized, the importance of an explicit geometric world model for enhancing the user's experience of a ubiquitous computing system has not been well-articulated. This paper introduces three scenarios which benefit from geometric awareness, examines three existing mechanisms for providing geometric knowledge, and then describes the EasyLiving Geometric Model. In particular, the focus is on improving the user experience of systems which are comprised of many independent devices and heterogeneous perception technologies.

Introduction

The goal of the EasyLiving research project[5] is the development of a prototype architecture and necessary technologies for intelligent environments. EasyLiving concentrates on applications where interactions with computing can be extended beyond the confines of the current desktop model. Such a computing system should maintain an awareness of its users, understand their physical and functional relationship to I/O devices, respond to voice and gesture commands, and be easily extended. This technology will, for instance, enable a home's resident to make a phone call by simply speaking his intentions from wherever he happens to be. It will allow a user to move from room to room while still maintaining an interactive session with the computer or a particular application. All these tasks require the coordination of many devices for computational activities involving both perception (Where is the user now?) and interaction ("Call Bob.")

In a space populated by many small, networked computing devices, several devices will typically have to work together to perform a particular task. Dynamically collecting a group of smart devices to enable an interaction or to perform a perceptual task requires a shared computational substrate that allows the devices to communicate bits which represent concepts in a shared ontology. For example, three devices might announce to each other over a wireless network: "I am a display device", "I am a DVD Player", and "I am an acoustic speaker." Once these capabilities are known to exist in the same place, the ability for a user to play a movie should be enabled. If a pair of headphones with appropriate capabilities were to enter the fray ("I'm a pair of headphones"), redirecting the sound output to them should become an available option.Where these options are displayed, how the user is informed of these options, and how much autonomy the system has remain open questions.

This paper introduces three scenarios which benefit from geometric awareness, examines three existing mechanisms for providing geometric knowledge, and then describes the EasyLiving Geometric Model. In particular, the focus is on improving the user experience of systems which are comprised of many independent devices and heterogeneous perception technologies.

Scenarios

Location-aware computing services have been proposed for many tasks, including providing driving directions, redirecting phone calls to the phone nearest the recipient, reminding the user of errands appropriate to his location, or even just turning off the lights when the user leaves a room. Most such systems assume a straightforward connection between the sensor providing the position information and the application. For example, the GPS position can be used along with a contact list and address book to deliver a reminder to pick up the dry cleaning when the user is near the store. The problem of performing location-based services in a more general framework, where there are multiple sensors and devices is much more challenging. A further complication is the incorporation of contextual information about the state of the physical world, beyond simply the user's position.

The following three scenarios describe how the user's experience of a ubiquitous computing system is improved by the use of "geometry-aware" system services. Geometry-awareness is distinguished from location awareness in two ways:

- heterogeneous perception technologies, e.g cameras, GPS, and beacons
- an understanding of physical relationships between things in the world, e.g. walls block users visibility

1. Contact with Context: Physical parameters for User Interfaces

Consider the task of contacting someone who is at work, such as to remind him of an important meeting. Which of the many devices in his office (displays, phone, pager, cell phone, computer speaker, stereo, PDA) should be used? One approach would be to use some fixed preference-based scheme, flashing the screen, ringing his phone, and finally paging him. However, there is no point in using a visual signal if he is not looking at the screen, or ringing his phone if he is wearing headphones. A better approach is to understand the location of the person, his physical relationship to devices which are around him, and the various consequences of the current state of the world. In this example, by examining the set of devices which are near the user (phone, pager, screen, speaker), examining the state of the world (facing away from the screen, pager inside briefcase, currently using phone), the system service which delivers the message can select the remaining option (speaker) to get the message to the user in the most expeditious manner.

The message-delivery example above illustrates the ability of geometric knowledge to provide physical parameters for the user interface. While geometric knowledge alone is insufficient to select the best device for a given interaction (context and other world knowledge are helpful), it is a necessary component for reaching the ideal decision.

2. Device Aggregation: Simplified device control

Currently, the PC is the integration point for a cluster of I/O devices which provide the majority of computer/information services. The wired mechanical connection of all the devices (processor, hard drive, display, speakers, mouse, keyboard, etc.) implies that this cluster of devices is intended to work together. What happens, though, when the devices lose this mechanical connection, as is the case with ubiquitous computing? If each device is an independent entity which connects to the network, some system infrastructure must exist to naturally pull together disparate elements to form a usable aggregation of devices.

Imagine your living room, equipped with a panoply of devices, including a couple screens, a wireless keyboard/remote, a sound system, a camera, etc. When you want to initiate a web browsing session, you would have to switch some screen to an appropriate mode using a specialized remote, manually direct the keyboard output to that screen, login as yourself (so you’ll have your cookies, preferences, etc.), tell the screen to enter web browsing mode, and manually redirect the audio to come from the PC. However, with geometric awareness, picking up the keyboard could implicitly log you in as yourself, and bring up an appropriate UI on a screen which you can see from your current location, and set up all the other parameters appropriate for your session. Additionally, the ability to dynamically change devices, such as having your session to follow you as you move, is enabled as long as exists some sensor which can provide the current state into a geometric model. Without geometric awareness, all device-to-task coupling must be performed manually, a task whose complexity will grow exponentially with the number of interconnected, networked smart devices.

3. Let there be light: Shared world model

Consider the simple task of turning on a light in an intelligent environment. Here are seven ways this task might be completed:

- Manual: Flip a wall switch.
- Traditional GUI: Use a dialog box with list of lights and buttons for on/off.
- Physically-enhanced GUI: Select from a map of house/room with lamp indicators.
- Direct Speech: Say “Turn on the living room lights.”
- Gesture: Make a funny gesture, observed by a camera, indicating a need for light.
- Indirect Speech: Say “I could use more light.”
- Implicit Request: Sit down at comfy chair while holding a book.

Some of these examples can be handled by existing technology and do not require any geometric awareness, or, in the case of “Manual”, even any computing technology at all. However, if the user and the computing system share some understanding of the physical world the ability to support a wider range of interactions becomes possible.

The “Physically-enhanced GUI” and “Direct Speech” use the least geometric knowledge. They only assume some simple shared map and nomenclature in order to enable requests. This requires the user to understand the building she is in, and the appropriate names of the various locations. Note that some tasks remain tough in this paradigm: how does one precisely name and differentiate all the various lights in one’s home? “Gesture” and “Indirect Speech” can use geometric knowledge to turn on the lights with more fidelity. Presumably, the user wants light to appear where she currently is located. By having a model of the location of the lights, and a perception system to track the user, the request can be more accurately fulfilled. The increase in capability of these interactions is directly related to the fidelity of the shared metaphor possessed by the computer and the user. The greater the shared understanding, the more robust the interactions can be.

Finally, notice that only the last option ("Implicit Request") does not involve direct explicit action by the user to request the system to turn on a light. The goal of geometric awareness is not just to provide automatic actions, but more importantly, to improve the shared model between user and computer, resulting in a more intuitive and natural user interface. We expect people to understand the consequences of their actions, and to interpret our implicit requests - why shouldn't we expect the same of our computers?

Three technology-driven notions of geometry

The above scenarios all describe possible ways in which geometric awareness can be used to support better situated interaction in ubiquitous computing. However, most existing systems which provide location information typically couple the mechanism for gaining position information directly to the services which are provided. In a more general ubiquitous computing environment, perception technology would be largely independent from the service - it doesn't matter how the computer knows the relationship between the person and the screen, only that the UI be able take this information into account when displaying information.

This section describes three extant paradigms for providing and utilizing geometric information in ubiquitous computing systems.

1. Latitude, Longitude, Elevation

Many proposals for location-aware systems rely heavily on position information coming directly from a GPS receiver, providing location with a resolution of approximately 30m. For some applications, differential GPS can be used to reduce the error to the sub-meter level; however, this requires significantly more hardware to achieve and constrains the system to operate within range of a differential-corrections transmitter[2]. Typical scenarios supported by such systems include using location to determine driving directions, deliver reminders based on the user's location[3], and record scientific data tagged with location information[4].

This notion of geometry is useful only in outdoor situations that are outside of major urban centers. In cities, the tall buildings can frequently obscure satellite visibility, much as when the antenna is indoors. While this does not preclude the usefulness of GPSs (and latitude/longitude measurements in general) for some scenarios, it implies other technologies for obtaining position information are needed. The above scenarios, for example, would all be impossible to achieve using only GPS as a positioning technology.

2. Beacons, Badges, Transceivers, & Tags

An alternative location-determining technology proposed primarily for indoor use is Active Badges. These systems can provide information about which room a particular tag is in[6], or even the particular position of the tag inside the room[7]. In general, these systems consist of RF, IR or ultrasonic transceivers which can determine the presence and perhaps location of small (usually powered) tags which are attached to objects of interest in the world such as people, phones, printers, computers, etc. These systems represent geometry as a location in a single coordinate frame, such as a map of the building. They require installation of a large number of transceivers throughout the building, and assume that all information of interest will be directly expressed in the single geometric model.

While a complete system to perform Device Aggregation scenario described above could be built by attaching tags to the devices (display, keyboard, speakers, user, etc.) and installing beacons in the rooms, there are two drawbacks. First, all items of interest must be tagged. If you are not wearing your badge (e.g. at home right after getting out of bed) then no position-based services are available to you. Secondly, and more significantly, if other positioning technologies are available, they cannot be readily integrated. Active badge systems are useful for providing positional information, much as is GPS, but current systems lack a general geometric model for expressing arbitrary geometric information.

3. Network Address != Physical Geometry

To avoid the perils perception altogether, one can assume that network or data connectivity is equivalent to co-location. This implies that if two devices can communicate directly (by RF, IR or other "local" transmission method), they are co-located. However, RF transmission (not to mention physical network protocols) can easily span rooms, floors or even buildings. Without some more precise model of geometry, this type of assumption will result in a excessively large set of potentially available devices, many of which may not actually be available or usable for any particular task. All devices are on the network (for some suitably comprehensive definition of network), and yet, the correct selection depends on knowing the physical relationship between the user and these devices. Worse yet, relying upon this alone can destroy the shared metaphor perceived by the user. For example, if a new reading lamp is plugged into an network/power outlet in the kitchen, but is routed around the corner into the den, asking for “lights in the den” to come on will not achieve the expected result. The user could be expected to manually configure the physical location of all devices, but this is a non-negligible burden. Additionally, consider the problem of any RF device relying on connectivity to determine location. If an RF remote control could control multiple devices in the same house, which device should it be activated when a button is pressed? If both devices can receive the signal, then either the user must resort to manual selection or confusion results.

The EasyLiving Geometric Model

The EasyLiving Geometric (EZLGM) model provide a general geometric service for ubiquitous computing, focussing on in-home or in-office tasks in which there are myriad I/O, perception, and computing devices supporting multiple users. The EZLGM is designed to work with multiple perception technologies and abstract the application and its user interface away from perception.

The base item in the EZLGM is an entity. An entity may represent an object or location in the physical world. Measurements are used to define geometric relationships between entities. In particular, a measurement describes the position and orientation of one entity’s coordinate frame, expressed in another entity’s coordinate frame. For two measurements to involve the same frame, both measurements must have been made by sensors with an implicit understanding of the origin of the frame on the entity; in other words, both must use a particular point and orientation on the entity when making an observation. Since objects in the physical world have some physical extent, this can also be expressed in the geometric model. If one physical object has different components which can be independently measured (e.g. a laptop with both screen and keyboard), then it could be represented as two entities.

Once a set of measurements has been provided to the geometric model the model can be queried for the relationships between entities’ frames. The measurements describe an undirected graph, which each vertex as an the frame of an entity, and each edge a description of the (invertible) geometric relationship (including an uncertainty estimate) between the coordinate frames. If at least one path exists between two frames, then the graph can be processed to produce a single geometric relationship between the frames. The replies to these queries are based on previously provided measurements. Since a particular queried relationship may not have been previously directly measured, the response is frequently involves the combination of multiple measurements; uncertainty information is used to accurately merge multiple redundant measurements as needed. When querying about the relationships between entities, it is frequently helpful to be able to refer to the extent of such an entity, such as the field-of-view of a display device.

To get a sense of how this model might be used, consider the Contact with Context scenario above. A process which wants to contact a user might query EZLGM for all devices which have service areas which intersect with the location of the user. It could then look at types and availability to determine the set of devices which might provide the messaging service, and further prune the list by considering the physical constraints, like visibility, etc., in order to reach a set of usable, available, and physically-appropriate devices. Visibility can be checked by examining all entities along the line of sight between the user and the device and ensuring none represent something which would physical block the view. Then, by consulting the users preferences, or by using internal heuristics and other context information, the signal could be sent via the appropriate mechanism. The geometric model provided a way for storing both the devices that could be used, and for aiding in the determination of the appropriate device. Note that no part of this example required any reference to the perception method which provided information about position: it could have been performed via cameras, a badge system, etc.

Conclusions

This paper has described three important scenarios for situated interaction in ubiquitous computing systems, “Contact with Context”, “ Device Aggregation” and “Extensible Computing.” Through those scenarios, three primary benefits of geometric models have been introduced:

- **Physical parameters for UI’s:** Device selection and control is performed with knowledge of physical context.
- **Simplified device control:** Device aggregation for a task is performed without requiring step-by-step user action.
- **Shared Metaphor:** User experience is simplified through a common understanding of the physical world shared by system and user

Existing systems which are location-aware provide applications which are tightly coupled to the perception mechanism. For location-awareness to be more generally applicable, abstracting the use from the gathering of such knowledge is essential. The EasyLiving Geometric Model is a first step towards a general sensor abstraction layer.

While geometric knowledge remains challenging to gather, represent, and provide, its inclusion will significantly improve user experience of ubiquitous computing systems

References

1. Dey, A. et al. "A Context-Based Infrastructure for Smart Environments", Managing Interactions in Smart Environments, Springer-Verlag, 1999, pp. 114-128.
2. "DGPS FAQ", http://www.navcen.uscg.mil/faq/dgpsfaq.htm
3. Rhodes, B. "The Wearable Remembrance Agent: A system for augmented memory", Personal Technologies Journal Special Issue on Wearable Computing, Personal Technologies, 1997, Vol. 1, pp. 218-224.
4. Ryan, N.S. et al., "Enhanced Reality Fieldwork: the Context-aware Archaeological Assistant", in V. Gaffney, M. van Leusen and S. Exxon (eds.) Computer Applications in Archaeology 1997, 1998.
5. Shafer, S. et al."The New EasyLiving Project at Microsoft Research", *Proceedings of the 1998 DARPA / NIST Smart Spaces Workshop*, July 1998, pp.127-130.
6. Want, R. et al. "The Active Badge Location System." ACM Transactions on Information Systems, 10(1):91-102, January 1992.
7. Ward, A. et al. "A New Location Technique for the Active Office", IEEE Personal Communications, Vol. 4, No. 5, October 1997, pp. 42-47
8. Weiser, M. "Ubiquitous Computing", http://www.ubiq.com/hypertext/weiser/UbiHome.html