Project Summary

Information systems are now mobile, wearable, multimodal, real-time, scalable from workstations to desktops to notebooks to palmtops to cellular, collaborative and ubiquitous. Network access is becoming increasingly important as a part of the computing infrastructure. Unfortunately, for the near term, unlimited bandwidth, anytime access to a network is not feasible, particularly in large urban environments where interference, occlusion and collision are ongoing problems. To alleviate this, we are developing a set of context-aware Autonomous Information Retrieval (AIR) pods. These are small, hardened low-cost computers that require only electric power. AIR pods have sufficient local storage to hold relatively static information and are equipped with two wireless interfaces: one short-range, unlicensed high-speed interface such as IEEE 802.11 and one long-range, low-speed interface such as CDPD or other Wide Area Network connection which may be intermittent. AIR pods can be either stationary or mobile. Stationary AIR pods can be attached to lamp posts and traffic lights, hidden in lighted store signs or in subway stations. Mobile AIR pods can be attached to delivery trucks, postal service vehicles, buses, police cruisers, taxis or other vehicles that roam city streets. AIR pods are also small enough to be carried in backpacks.

We propose to develop and implement the prototype hardware and software AIR idea, and explore what kinds of infrastructure support are necessary to make these devices a key component of network applications. Research issues addressed will be cooperative data sharing, resource scheduling and anticipatory caching, message propagation and wide-area resource discovery.

The proposed mobile networking infrastructure will be tested in a demanding set of context-aware mobile research projects. The first project is a wearable augmented reality system that allows outdoor users to tour a campus interactively. A second related project allows indoor users to collaborate with those outside over wireless networks. The third project is a mobile robot sensing system that can autonomously explore the campus and create rich 3D, texture mapped, site models. All of these applications need to interact with host computers through limited network access, and we propose to optimize this interaction over bandwidth, devices, and locality using context-aware wireless networks.

1 Infrastructure for Context-Aware Wireless Network Applications

In the years to come we envision mobile users roaming through metropolitan areas enabled with pervasive computing and communications. They will use their cellular phones, PDAs, wearables, laptops or other personal mobile devices to run context-aware services. To support these services, there will be a large number of sensors monitoring or measuring entities for a certain context (such as temperature, motion, location, presence of people or physical objects, speed, "state" of a room or a user) and producing data. In addition, physical objects, with RF-tags or low-power, low-range network interfaces will beam information to the users in close proximity that query them. Context-aware [26, 4]. applications on the mobile devices will query for these sensors' data or other information as needed. Imagine a tourist walking through a major city like New York, overwhelmed by the size and scope of such an environment. However, a set of smart, Autonomous Information Repositories (AIR) placed strategically around the city, could provide access to local information sources and wider area connections. AIR pods can provide wireless communication to the tourist's own PDA/cell-phone/laptop connections seamlessly. In essence, the tourist never is out of touch with an online Michelin guide, but a guide that is much more capable, dynamic, and up-to-date.

Besides serving as information repositories, AIR pods can also serve as gateways to upload information from a tourist. Photos and videos taken at a site can be transmitted upstream at the site with time, date and location information stamped on them. Instant email postcards can be generated at the site by the tourist.

We envision building an intelligent network of AIR pods that can not only provide the services outlined above, but can also create and modify user models to prefetch and store relevant information for the tourist at accompanying local sites. The network can also serve as a feed-forward device to suggest where the tourist may go next on his tour, which can then supplement the user model with new information.

AIR pods need to be scalable to provide low bandwidth information to PDA's, and high bandwidth to more sophisticated "Touring Machines" (section 4.1). They also need to be able to accept data streams from mobile sensor units that may be working in their environs that create online 3-D models of the environment (section 5).

This proposal addresses the needs of providing context-aware information to mobile wireless users through the AIR mechanism.

From a network perspective, the city tour guide application offers a number of challenges. The video overlays and architectural renderings for each building or other destination can each consume several tens to hundreds of MB. This information, relatively static, needs to be supplemented by more rapidly changing information, for example on special events, transportation schedules, weather forecasts and possibly advertising. Current end systems, such as PDAs, have, at best, tens of MBs of available storage and thus are not suitable to store a whole city tour.

While our initial implementation will be focused on the city guide application described above, we envision a range of other location-based information services and intend to prototype a general system suitable for urban information access and distribution [29, 16, 11]. Examples of location-based services are notification systems, e.g., sending a reminder about the special offers when approaching a store or bus schedules when approaching a station.

Mobile users not only want to retrieve data, but also provide information to the network. Examples include uploading digital images while traveling, GPS-equipped cars sending their speed to the local traffic report server, or "spatialized post-it note" [7, 32], where users store location-specific information.

We envision that these applications are supported by autonomous information repositories (AIR) pods, small, hardened low-cost computers that require only electric power. AIR pods have sufficient local storage to hold relatively static information, such as the video clips for attractions in the vicinity. Each AIR pod is also equipped with two wireless interfaces, one short-range, unlicensed high-speed interface such as IEEE 802.11 and one long-range, low-speed interface such as CDPD or any Wide Area Network (WAN) connection which may be intermittent. AIR pods can be either stationary or mobile. Stationary AIR pods could be attached to lamp posts and traffic lights, hidden in lighted store signs or in subway stations. Mobile AIR pods can be attached to delivery trucks, postal service vehicles[1], buses, police cruisers, taxis or other vehicles that roam city streets. AIR pods are also small enough to be carried in backpacks.

Fig. 1 describes wireless information architectures along the dimensions of control and use of network infrastructure, either wired or wireless with fixed base stations. Below we outline the issues involved in creating an intelligent network of AIRs, discuss algorithms to allow AIR pods to serve a mobile user, and discuss 2 mobile applications already in place on the Columbia campus that can serve as testbeds for the AIR concept.

---

[1] Here, the US Postal Service and city busses can take on a new role. These vehicles "empty" the stored encrypted messages from AIR pods and delivers them to their recipients and could also refresh the content stored in AIR pods.

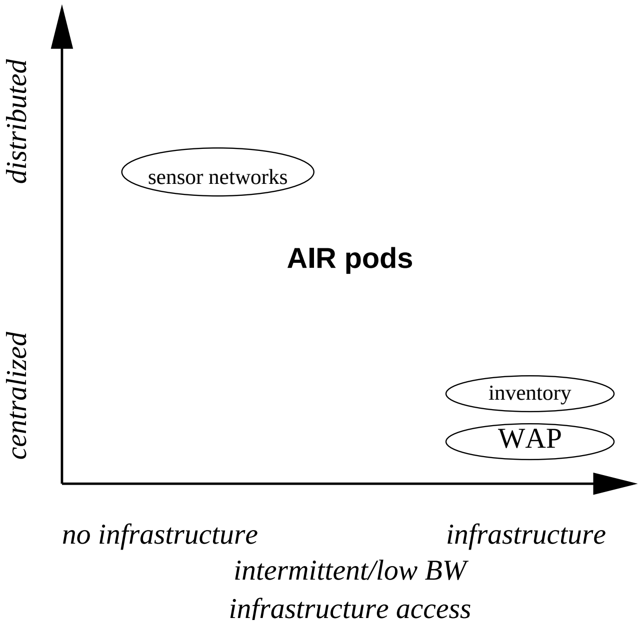

Figure 1: Relation of project to other work

2 Technical Problems

The applications noted above are characterized by high data volumes. For example, high-quality video with MPEG compression consumes at least 6 Mb/s, while compressed megapixel JPEG images can take up several MB.

All of these examples can be implemented using relatively standard Internet technology, but these implementations rely on high-speed continuous network access and full-function end systems with large storage capacity. Despite predictions to the contrary, high-speed wireless data access outside campus areas and possibly airports is not likely to appear in the next decade [6]. Current first and second-generation wireless systems, such as CDPD, CDMA or GSM, have wireless capacities of about 10 kb/s. Third-generation wireless systems such as UMTS are likely to offer at most 384 kb/s across their coverage areas, with higher speeds possible close to base stations[2]. Second-generation Ricochet networks are limited to 128 kb/s and available only in major metropolitan areas.

However, spectrum space in frequency regions where radio waves can traverse walls or travel significant distances is severely limited and expensive. (Spectrum auctions routinely value spectrum at about $1,000 per potential customer, without a single transmission tower having been built.) Thus, in addition to scarce, licensed wireless bandwidth will remain expensive. For example, CDPD data transfers cost about $0.1/kB.

2.1 Cooperative Data Sharing

AIR pods and users can share data amongst themselves. This allows popular data to spread without any infrastructure. Example of such data includes newspaper articles, music and video content (presumably legal), and maps. Unlike for normal web retrieval, mobile users do not have a whole web site to browse. Thus, instead of clicking on a link on the newspaper home page, the mobile terminal announces to the radio neighborhood that the owner is looking for news on politics. We envision that mobile terminals have built-in micro search engines that use meta-information embedded in content or textual information to check whether they can supply this information.

The system exploits the high locality of information access within a geographic area, especially, in smart spaces, where sensors and physical objects are beaming information to the mobile users. Another characteristic is that people on the move are likely to be more flexible in their information tastes, media format or quality as well as in the information accuracy.

We have done initial simulation experiments emulating the behavior for a single data item [21]. We simulated

---

[2] Technologies likely to appear in the next decade, such as GPRS, are still limited to 15 to 30 kb/s in realistic use scenarios.

the distribution of data within an area roughly equivalent to a few city blocks, representing for example midtown Manhattan. Manhattan has a population density of 4,434 people/km^2. In our models, we assume that between zero and 50 people are interested in a particular data item. This represents a rough estimate based on the popularity of the *New York Times*, read by about 8% of the New York City population and the assumption that about one tenth of those are walking around during the morning or evening rush hour. In the random walk model, information holders walk a random distance in a single direction, then stop and change course. We define the *query interval* as the time that lapses between two consecutive queries sent by the same user. Fig. 2 show how long it takes until all those interested in the data have received it, while Fig. 3 indicates the average delay until a user finds the information desired.

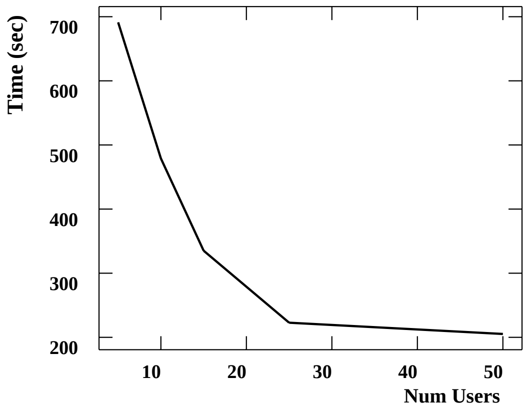

Figure 2: Time until full propagation, random walk model (60 s query interval)

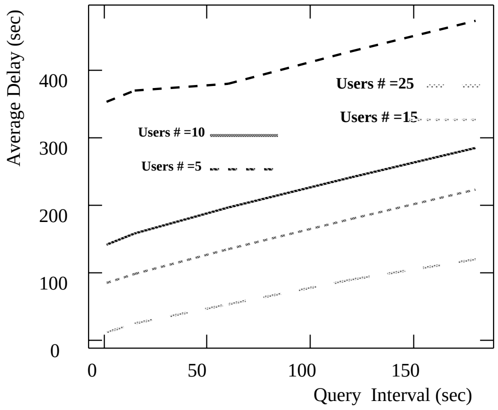

Figure 3: Average delay to acquire data, random walk model

We also simulated a model reflecting a metropolitan subway. We assume that passengers holding the datum under consideration arrive at the subway station as a Poisson process, with an arrival interval of between one and three minutes. Trains arrive every five minutes and stop for 45 seconds, making ten stops and taking between 168 and 210 s between stops. Trips are between two and six stops long. Data is exchanged while passengers are waiting on the platform and with other passengers while riding the train. Fig. 4 shows the fraction of passengers that have obtained the desired data at the end of their commute, while Fig. 5 depicts the average delay.

As part of this proposal, we plan to implement data sharing using AIR pods and standard PDAs equipped with 802.11 cards. This requires suitable estimation of content categories, development of a small-footprint search engine and a web-compatible protocol for announcement and retrieval. Issues of content authentication and access control also need to be investigated.

Since the residence time of a user or other AIR pod within transmission range an AIR pod is unknown and may be short and since transmission bandwidth needs to be shared between users in range of the AIR pod, we plan a system

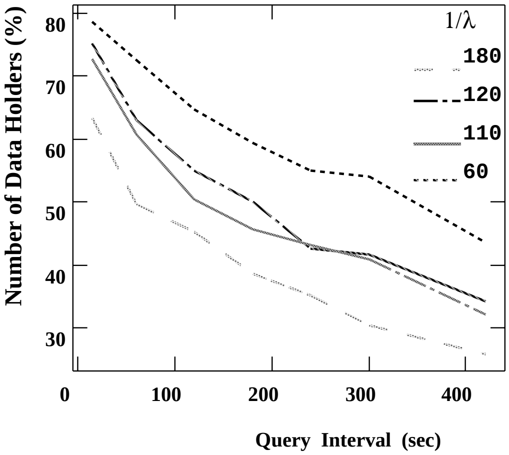

Figure 4: Percentage of data holders in the subway model

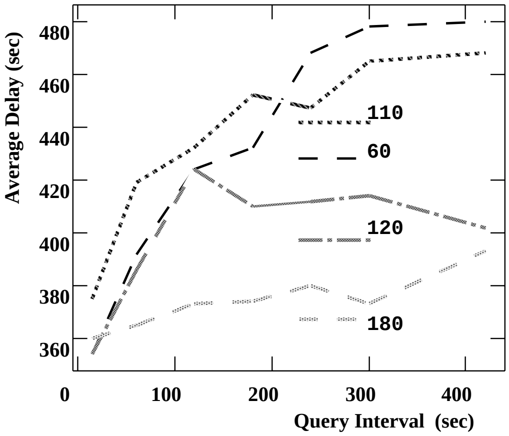

Figure 5: Average delay for data propagation in the subway model

of progressive (or layered) delivery, where content is delivered in stages, with the smallest resolution being available first. A variation of the pyramid or skyscraper broadcasting mechanism [14] may also be applicable here, as the radio technologies naturally support multicast delivery. It appears interesting to combine these techniques with layered coding [2, 27].

2.2 Resource Scheduling and Anticipatory Caching

For the city guide application, most of the content will be fairly static, as long as it fits into local storage. Given the current cost of low-speed wireless access, uploading videos is not likely to be feasible. However, if the AIR pod is connected to the Internet via a lower-cost connection, such as a dial-up modem, it may be feasible to update the content before a user reaches the radio range of the AIR pod. For example, if a Norwegian tourist group is walking up Fifth Avenue, the AIR pods in the anticipated path of the tourist can retrieve the Norwegian-language version of the tour guide, even though it is not efficient to store it permanently on each AIR pod, given the relatively low frequency of Norwegian tourists. However, the AIR pods do not know with certainty which nodes the tourists will be visiting next. It has some notion of direction and speed, but it is quite possible that, say, the tour group will stop in a cafe or turn into a side street. Thus, cache expiration strategies need to be devised that optimize the overall utility of limited storage space.

There has been much work on web prefetching for latency reduction (e.g., [18]), particularly for low bandwidth links, but this work does not take into account the use of geographic information to predict retrieval patterns or the different access patterns and needs of mobile users. In addition, it reduces latency at the expense of higher data volume, a solution that works for circuit-switched connections, but which is not appropriate for shared wireless bands in which users may pay by the KB.

There have been some initial initial efforts to make use of geographic info [5]. We will extend these to take into account both content information and geographical trajectory information, based in part on knowledge about our application. For example, predictions can be based on the interaction history of other users, as well as on the linking structure of the material that the user is currently exploring.

Unlike earlier work [15], we plan to look at global optimization of utility, not just greedy per-user optimality.

2.3 Message Propagation

In addition to using their low-bandwidth wide-area wireless channel, stationary AIR pods can also transmit data to the Internet by using the services of mobile AIR pods. Mobile AIR pods that anticipate reaching an Internet access point "pick up" messages from stationary AIR pods, acting as third-party data carriers. Messages can be as simple as email messages or as large as digital images or movie clips. The messages are encrypted so that carriers can only see the destination address. Absolute reliability is not required, as we anticipate that individual messages are not critical (e.g., for many measurement applications) or the data source eventually can drop off the message itself. Thus, inter-AIR pod message propagation accelerates the delivery of messages.

We plan to develop a delivery protocol and mechanism that replicates messages to several carriers, based on updated estimates of delivery probability and urgency. The AIR pod also provides the carrier with an estimated life time for the message, based on its computation of when a copy will likely have reached an Internet access point.

It may also be possible to have the carrier or Internet access point deliver a brief acknowledgement to the AIR pod that the message has reached its destination, terminating retransmission attempts.

Multiple-hop routing may also be possible. For example, a carrier that recognizes that it is not making sufficient progress towards the nearest Internet access point may transfer its messages to another AIR pod that is passing by.

2.4 Wide-Area Resource Discovery

As we described previously, context-aware services query for data as needed. Queries and data distribution have geographic locality of reference. A location-based hierarchy for the information storage and the data filtering may reduce substantially the communication cost. For our tourist guide example, visitors may want to explore the physical neighborhood, asking questions such as "what other monuments are around here?" or "are there any museums in this part of the city?".

We had earlier studied the problem of wide-area service discovery for locating the best gateway from the Internet to the traditional telephone network [25, 24]. The model is based on the notion of aggregation, where each "region" exports a summary of its services to its neighbors (Fig. 6). For example, all Internet telephony gateways (GW) export tuples listing area code ranges and their cost of reaching them, depending on whether it is a local or toll call.

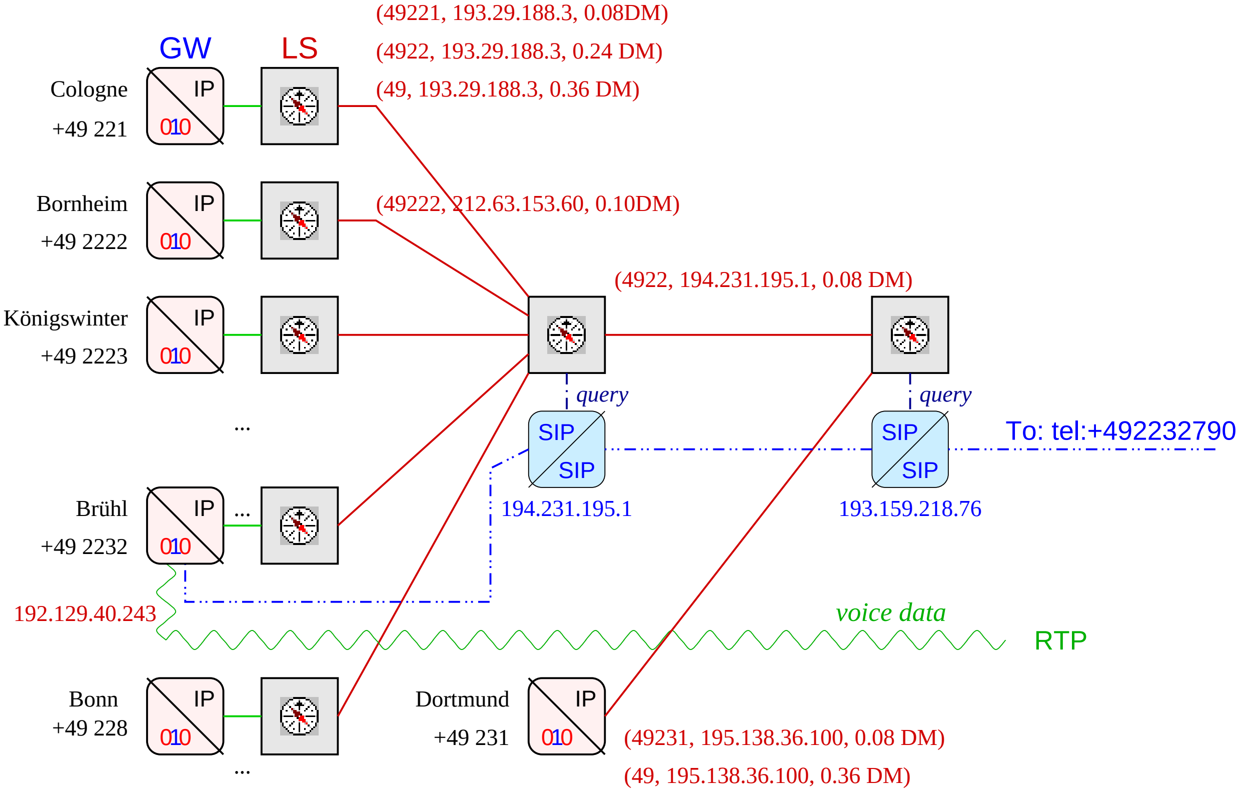

Figure 6: Example of gateways (GW) aggregating summaries of their local services (LS) to neighbor gateways to enable service discovering in the wide area network.

This architecture allows aggregation of information, so that only summary information (e.g., “The gateways I represent can reach all of Europe for $0.20.”) needs to be exchanged across wider areas. Since local information may change rapidly, this prevents having to flood the Internet with these updates or keep everything in a central server, prone to local overload or failures. The current protocol design, TRIP, uses BGP to distribute this information, but associations (“peering”) and hierarchies need to be configured manually.

We plan to generalize this architecture so that it is suitable for other types of information, such as sensors data (e.g., traffic information cited earlier). Specifically, we would like to design an infrastructure of hosts attached to data repositories. The hosts act as application-based routers to locate the data for the applications. The hosts gather the raw data produced by the sensors or other providers and store them in the repositories together with supporting XML-style meta-data or attributes. Queries for the data are expressed using these attributes. By collecting data that either correspond to measurements from more than one sensor or to a longer monitoring period or to more layers (in the case of multimedia layered coding), the information certainty (accuracy) may increase. Moreover, applications have different requirements for information certainty and may trade it for a faster response. In this infastructure, the hosts learn about their stored contents and form hierarchies in a self-organizing manner. A location-based hierarchy (as in the previous example) can be the ”core” hierarchy of the infrastructure. We plan to support both single-root trees as well as meshes of trees, where the top level nodes are peers. We want to design a scalable system that reduces the latency to locate the relevant data.

2.5 Related Work

The DataMan project [16, 20] envisions a similar setting. Our system differs from this work in the design of the infrastructure. As we mentioned previously, we consider hosts that in a self-organizing manner form caching hierarchies

Figure 7: The Columbia Touring Machine. Mobile outdoor user wearing backpack and see-through head-worn display (right) sees strategic information overlaid on the real world (left). Complementary material is presented on a hand-held computer (lower left).

and act as application-based routers to locate the relevant data for the queries sent by the context-aware services on the mobile devices.

The WINLAB Infostation [9, 15] is also based on the assumption that high-speed wireless network access is less than ubiquitous. Infostations are stationary devices with Internet access and a short-range radio transmitter.

In addition, we extend the network model to allow the propagation of information from user to user and AIR pod to AIR pod, not just Infostation to user. In our model, AIR pods can also be mobile, raising the issue of message propagation. Our hardware solution is readily deployable for related experiments as it relies on standard protocols and wireless networks.

3 Outcomes

As part of the project, we plan to architect and design a self-configuring AIR pod that only needs electric power. We anticipate using common PC/104 (90 by 96 mm) components that can be carried in vehicles and hidden in buildings, providing limited amounts of CompactFlash storage (currently, 64 to 340 MB, with at least 1 GB anticipated during the project), an IEEE 802.11 11 Mb/s interface and a CDPD, ARDIS or BSWD or other low-speed wireless interface. Mobile AIR pods are also equipped with a GPS receiver to estimate speed and position. We anticipate that each such data station costs about $1,500.

We propose to build a system that enables mobile users to cooperate and share data. In addition, we will design and implement the important components of a scalable infrastructure allows context-aware services to be made available to mobile users, while reducing latency.

4 Distributed Mobile and Wearable Augmented Reality Systems

As computers increase in power and decrease in size, new mobile and wearable computing applications are rapidly becoming feasible, promising users access to online resources always and everywhere. This new flexibility makes possible a new kind of application—one that exploits the user's surrounding context. Context-aware (and, especially, location-aware) computing [26, 4] allows us to link electronic data to actual physical locations, thereby augmenting the real world with a layer of virtual information. *Augmented reality* [3], in which displays overlay graphics, audio, and other modalities on the real world, is particularly well suited as a user interface (UI) for context-aware applications. Equipped with location and orientation sensors and with a model of the user's environment, a wearable computer can

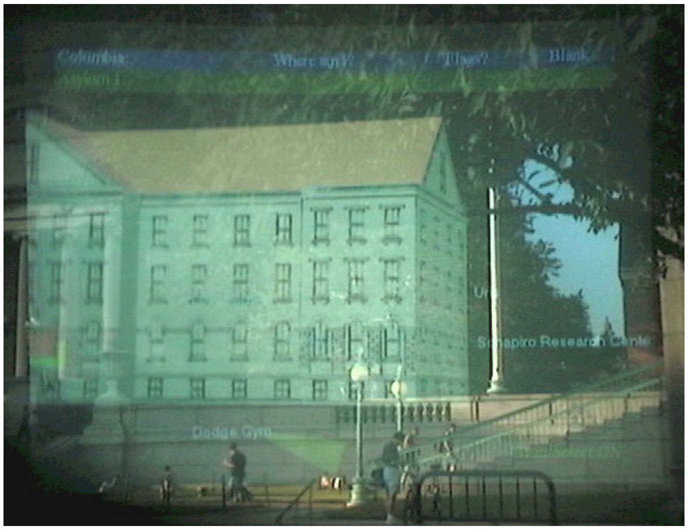

Figure 8: Model of demolished building, imaged through see-through head-worn display.

annotate the user's view of the physical world. Through optical or video-mixed see-through displays, the user views the electronic information in situ, attached to the physical world, and can interact with this virtual layer or even modify it. Thus, the world *becomes* the UI.

Wearable UIs alone will not be enough to fully exploit the potential of a world-wide layer of spatialized information. Many tasks may benefit from the use of physically large, stationary displays integrated into, or supported by, room surfaces and furniture. Such tasks include ones that involve collaboration among collocated users, such as authoring parts of the information space, or obtaining a shared overview of large remote portions of that space and the ways in which other users are interacting with it.

The proposed mobile networking infrastructure will be tested in a demanding set of context-aware mobile and wearable augmented reality research projects. Our projects stress networked interaction among multiple machines and users through a distributed virtual environment infrastructure [19, 13], and include both indoor and outdoor testbeds.

4.1 Outdoor Testbed

One of these testbeds is our "Touring Machine," a prototype wearable augmented reality system [8]. As shown at the right in Fig. 7, the Touring Machine's mobile user wears a color SVGA stereo, see-through, head-worn display, and holds a stylus-based computer with its own color VGA display. The user's head orientation is tracked with an inertial 3D tracker and their position is tracked by a centimeter-level RTK GPS system. This GPS system was obtained through previous NSF funding, replacing a far less accurate submeter differential GPS system we had used previously. Used in conjunction with the head orientation tracker, the RTK GPS allows spatial information, extracted from a campus database, to be overlaid on and registered with the mobile user's view of the real world as she walks around outside. The head-worn display is driven by a backpack that contains a PC with 3D graphics hardware, connected to our wired infrastructure through a radio network interface. While our current system is heavy and bulky, especially in comparison with the work of researchers who emphasize current wearability (e.g., [28, 32]), we are willing to sacrifice physical comfort for programming ease and computational power to better approximate the capabilities of future machines, especially tracked 3D graphics.

Our head-worn UI, shown at the left of Fig. 7 and in Figs. 8, and 9 (all photographed directly through our see-through head-worn display), consists of a world-stabilized part and a screen-stabilized part. World-stabilized items are visually registered with 3D physical locations, and displayed in the correct perspective for the user's viewpoint. They include labels of buildings, iconic flags representing important information that can be further examined by the user, and virtual representations of physical buildings (Fig. 8).

Screen-stabilized items are fixed to the display and are always visible; they include the menu bar at the top, and the cone-shaped pointer at the bottom. Head-worn display menus are controlled through a trackpad mounted on the hand-held computer. The cone-shaped pointer indicates the currently selected object. An object can be selected through several mechanisms: an approximation to gaze-directed selection that targets the closest object to the head-

Figure 9: Situated documentary, photographed through see-through head-worn display, shows suggested path leading between flags that represent parts of hypermedia documentary.

worn display's center that is within a bounding volume; direct selection of links presented in the head-worn screen menu or on the pen-based hand-held computer; selection based on positional proximity in the real world; and selection through program control.

The hand-held computer runs a variety of applications, such as the 3D campus map shown at the lower left of Fig. 7, whose display can be coordinated with the head-worn display. The hand-held computer also runs a standard web browser that is connected to a custom HTTP server, allowing our applications to feed web pages to the browser and respond to interactions with it [8].

Models of 3D infrastructure, such as buildings or underground tunnels, can be shown on the see-through head-worn display, registered with the physical world. For example, Fig. 8 shows a 19th century building that was demolished prior to the construction of Columbia's current campus, overlaid at its original location.

Working with with colleagues in Columbia's School of Journalism, we have used these capabilities to construct multimedia "situated documentaries" [12] that inform users about their surroundings by embedding a hypermedia documentary within the 3D physical environment that it describes. The iconic flags shown in Figs. 7 and 9 represent synchronized multimedia chunks of narrated imagery, sound, and video that can be selected for presentation by the user. Fig. 9, imaged from the top of a campus building, shows a suggested path for the user to follow. We are currently developing the next generation of this system, which will support collaboration among multiple mobile users, both collocated and dispersed across Columbia's campus [13].

4.2 Indoor Testbed

Users of our outdoor augmented reality system can collaborate with indoor users. To support the design of our situated documentaries, indoor users interact with different representations of the campus, ranging from a 2D map displayed on a hand-held computer, to a shared 3D virtual "table-top" model of the campus, presented to multiple users wearing head-worn displays. Fig. 10 shows this simplified 3D model of Columbia's campus, which is positioned so that its ground plane is roughly coplanar to a physical desk. Indoors users can interact with the model to create virtual objects and highlight and annotate real objects for outdoor users to see, and maintain histories of outdoor users' activities; in turn, outdoor users point out interesting objects and events for indoor users to view. Our infrastructure allows all applications to access databases that contain models of the physical and virtual environments, maintained internally in a relational format [13].

4.3 Proposed Work

In these projects, we need to transmit real-time information among outdoor and indoor mobile users and systems, including speech, gesture, multimedia, program data, and RTK GPS corrections. For example, users of our Touring

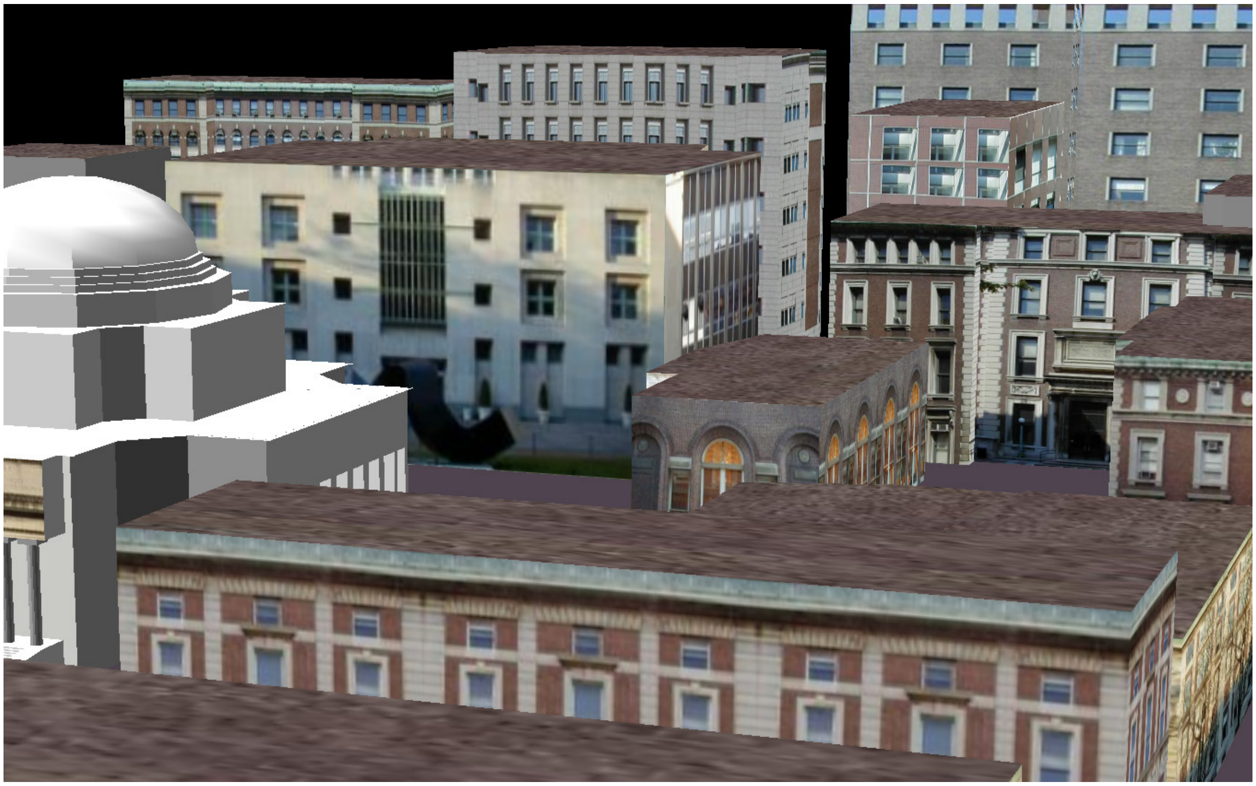

Figure 10: 3D model of Columbia's campus viewed by indoor users.

Machine view web pages that include dynamic media, such as MPEG videos, as part of our situated documentaries. We would like mobile users to be able to access this material when and where it is needed. We currently download all needed material in advance, with the mobile machine attached to the wired network. Unfortunately, this approach eliminates spontaneous presentation of unanticipated multimedia data that is too large to stream. The proposed network infrastructure would instead allow us to experiment with a variety of alternatives.

Here are some of the approaches that we propose to investigate:

- Information transfer among nearby mobile human and robot users can be effective when some users are better situated with regard to the wireless infrastructure than others. However, there may be times in which a relatively small change in a user's location may result in a significant improvement in their connectivity. If there is some way for the system to determine this, either from current connectivity data obtained from others or from historical data, how should it be communicated to a human user? One option would be to provide explicit verbal or visual instructions telling a user where they can move for better connectivity. An alternative might be to offer more subtle cues in the spirit of background "ambient displays" [17], such as gently highlighting areas into which it would be good to move when a user's current connectivity is inadequate

- While users of our situated documentaries are free to jump around among the available material, many parts are designed to be viewed in sequence and can be most conveniently viewed this way. Suppose that a user will encounter data in some suggested order, while following a pre-planned path (such as that of Fig. 9), perhaps with known branch points. Can we take advantage of this to determine when to transmit information to the user, based on predicted and current connectivity at each point in time along the path? What information would need to be provided by the application to the network infrastructure? Or might the network infrastructure be able to predict what the user will need and when it will be needed by matching patterns in her computer's data requests with historical data?

- How can the significant connectivity improvements that might be gained from small variations in a user's location be taken into account in the original design of a situated documentary? How might a pre-planned path be modified on the fly to satisfy the constraints imposed by the current network status while preserving the integrity of the information to be presented?

Our mobile and wearable augmented reality systems will provide an intensive test environment for the proposed network infrastructure. By addressing how mobile human–human and human–machine interaction, outdoors and indoors, can be better supported through improved mobile network infrastructure, we will develop prototype user interfaces that better address the needs of mobile users.

5 Mobile Sensing Applications

As wireless access becomes available, autonomous mobile sensors can become an important part of the computing infrastructure. Sensors can be placed in remote areas while maintaining a link to a host, possibly using an AIR pod as an intermediate waystation. Since sensors can be mobile, they can plan where to go, and in turn, the network can be modified according to these plans to provide support to the sensing operation.

We want to extend the idea of AIR pods to mobile sensing. There are three specific applications that we are currently exploring that can make use of this context-aware network infrastructure.

5.1 Automated 3-D Site Modeling

The first application is automated 3-D site modeling. Site models are used in a number of applications ranging from city planning, urban design, fire and police planning, virtual reality modeling and others. This modeling is done primarily by hand, and owing to the complexity of these environments, is extremely painstaking. The models built are often incomplete and updating them can be a serious problem.

As part of a previous NSF supported equipment grant (EIA-9729844) we have built a mobile robot system called **AVENUE: Autonomous Vehicle for Exploring and Navigating Urban Environments** that autonomously moves around an outdoor urban site and creates a photorealistic, geometrically correct 3-D model of that environment with limited human interaction [10, 30, 31, 23]. For a site modeling task, the robot is provided with a 2-D map of its environment. High level planning software is used to direct the robot to a number of different sensing sites where it can acquire imagery that is fused into photo-realistic (i.e texture mapped) 3-D models of the site [30]. The system must plan a path to each sensing site and then control the robot to reach that site. Positional accuracy is a paramount concern, since reconstructing the 3-D models requires precise registration among image and range scans from multiple acquisition sites.

The mobile robot base that we use for our experiments (see Figure 11b) is an *ATRV-2* model manufactured by Real World Interface, Inc. It is an all-terrain robot vehicle utilizing four-wheel drive and differential steering. Two GG24C Surveyor GPS+GLONASS receivers running in RTK/CPD (Real-time Kinematic/Carrier Phase Differential) mode provide us with positioning information. One of them is installed as a base station in the roof of a tall nearby building and provides, over a radio link, differential corrections to the other one, on the robot. With a sufficient number of satellites visible (usually 7 or 8) this setup gives us 1Hz position updates accurate to a few centimeters. A color CCD camera is affixed to a pan-tilt unit (PTU) mounted in the front of the robot, and is connected to a frame grabber installed in the on-board computer. We have also mounted a Cyra Technologies Cyrax 2400 laser range scanner on a custom-built platform attached to the robot. The scanner uses an eye-safe class II laser and provides variable resolution scans in the range of up to 100 meters. These two sensors are the primary acquisition devices for the site modeling task.

The modeling process to create 3-D models of outdoor sites with *AVENUE* is depicted in Figure 11a and is described in detail in [1, 22, 30, 31]. *AVENUE* roams around an environment (in our case the Columbia campus as shown in Figure 11c), and acquires 2-D images and 3-D range scans of buildings. It then builds these into correct, coherent, photo-realistic 3-D models that can be integrated with an existing 2-D campus map. The nature of this sensing environment (cluttered, with high occlusion for vision sensors) precludes exhaustive sensing of the environment and necessitates a planning function to allow the robot to build the models efficiently. Our method plans the next sensor viewpoint so that each additional sensing operation recovers object surface that has not yet been modeled and attempts to reduce the number of sensing operations to recover a model. Systems without planning tend to use human interaction or overly large data sets with significant overlap between them. Given large data set sizes and long image acquisition times, reducing the number of views while providing full coverage of the scene is a major goal. The planning algorithm [22, 23] is able to reason about scene occlusion, sensor imaging parameters and sensor placement

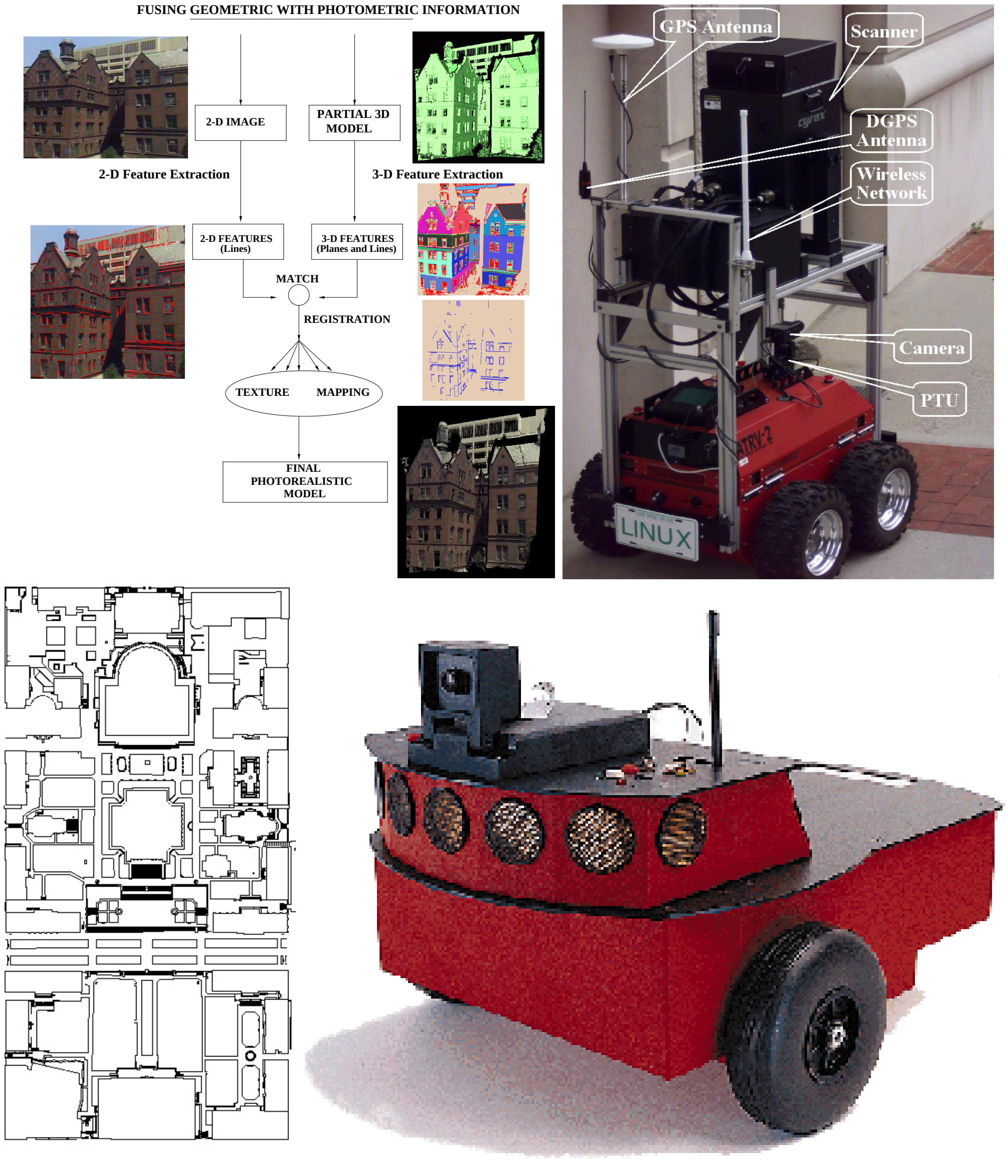

Figure 11: A) Method for photorealistic 3–D modeling. B) Mobile Scanning AVENUE robot. C) Map of the Columbia campus where the wireless network is being established. D) Pioneer mini-robots, equipped with AIR pods can act as mobile agents for data sharing in Ad-Hoc networks

Figure 12: a) Image of the scene, b) 3D model of the scene and c) Image texture-mapped on 3D model after the registration.

constraints to compute valid, occlusion-free viewpoints for the robot's sensors. Once an unoccluded sensor position for the surface has been determined, a new image is acquired, modeled, and integrated with the model. Thus, the method is target-driven, and as modeling process proceeds, an occlusion-free sensor viewpoint will always be found for regions that require additional sensing, if one exists. The sensor parameters planned for include field of view, resolution, and standoff distance. Figure 12 shows a 2-D image of a building on the Columbia campus, a 3-D model of the building acquired from the mobile robot, and the texture mapped model which integrates the 2-D imagery. Figure 13 shows 3 segmented range scans of another building on campus, the registered depth maps, volumetric sweeps of each scan, the integrated 3-D volumetric model of the 3 scans, and a texture mapped rendition of the integrated model. Colored regions denote planar segmentations of the range data.

Once a new sensing position is determined, it remains to navigate the robot to that position. The mobile robot needs to maintain communication and connection to host computers that contain 2-D maps, the current 3-D models and other geographical/positional information that can not be stored locally on the mobile agent. The mobile agent can be seen as a distributed sensor that can be sent to regions to gather data, but this data needs to be transmitted back to the home station for integration into a global model. Data will also have to be sent to the mobile agent for control functions and updating of onboard local map information. The mobile agent will need "directions" to find its way to a planned sensing site. The glue for all this is the combination of GPS and mobile network access as described previously. The GPS serves as a reference for positioning information, defining network topology and access points, and allows different communication modes to be established. The AIR pods can form an intelligent network that can place navigation data in locations where the robot can receive it and knowledge of AIR pod locations will allow the robot to navigate to a live network connection area to upload sensor data.

An important area of research in this proposal is prestaged delivery of data at points in the network topology. Consider a site model for a large area such as a city. Models can be maintained at many levels of detail, from 3-D geometric CAD models, texture-mapped images overlaid on 3-D models, interior space diagrams, and at even lower levels wiring and piping access inside the walls of the structures. A mobile agent can be used to update, enhance and refine these models in an autonomous fashion. It clearly cannot store the entire site model locally, but will need portions of it to navigate and report back sensory data. Hence, we can download relevant model portions to the network nodes whose topology is consistent with the robot's current position and its future positions since we are controlling the robot, and are receiving GPS feedback on its position. The central controller can maintain position based routing tables for the robot, and pre-stage ontime deliveries to the robot as it needs them for new navigation and control commands based upon its current and predicted location the cache-sharing techniques described in section 2.2.

5.2 Intelligent Mapping of Network Topology

The mobile base is capable of navigating autonomously, and through its GPS connections, it can maintain locality information. This mobile sensor base can be used to efficiently map out a network topology, finding areas of good, bad and no wireless connectivity. Effectively, the mobile base can create a mapping of bandwidth versus location, and this mapping can be effectively used by resource allocation algorithms (section 2.2) to optimize content placed in AIR pods, some of which may themselves be mobile. As users become GPS equipped (such as in the Touring Machine

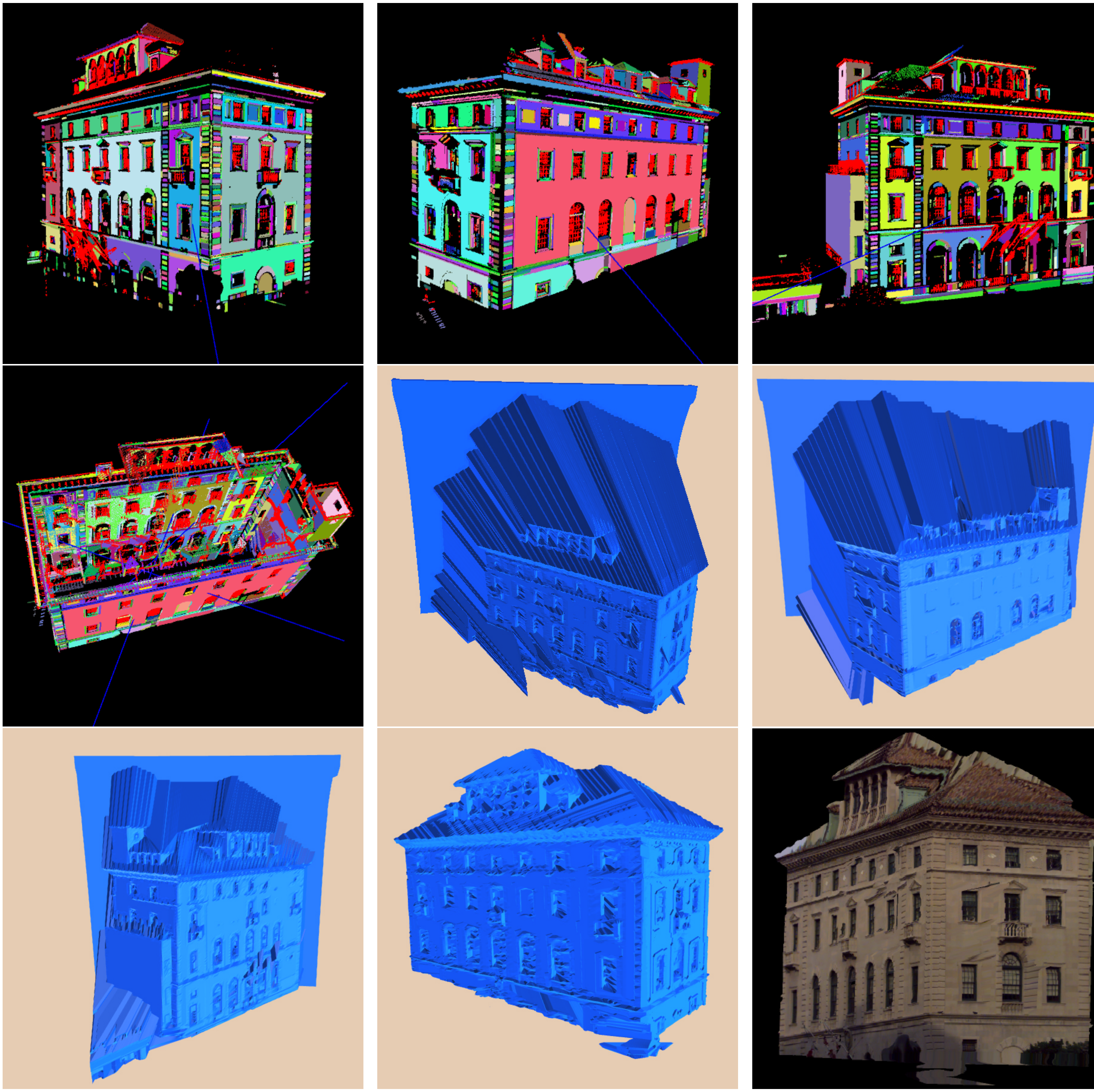

Figure 13: Model Building Process: a) First segmented depth map, b) Second segmented depth map, c) Third segmented depth map, d) Registered depth maps, e) First volumetric sweep, f) Second volumetric sweep, g) Third volumetric sweep, h) Composite solid model generated by the intersection of the three sweeps, i) Texture-mapped composite solid model.

application, their locations and network topologies can be easily mapped onto each other, allowing a predictive and efficient way to make sure data is available to mobile users.

5.3 Data Sharing in Ad-Hoc networks

We will use our mobile distributed sensors to explore the paradigm of data sharing in Ad-Hoc networks. We will use an existing cadre of small, dedicated explorer robots (Pioneer I mini-robots, figure 11d)) that maintain a dedicated radio modem link to the mobile base. These robots can be used for small sensing tasks as they are equipped with cameras, and can report back to the larger mobile base unit which maintains wider scale connectivity. Thus the mobile base becomes an intelligent router on its own, sending data between the mini-robots, and also upstream to the central controller.

References

[1] Peter K. Allen, Michael K. Reed, and Ioannis Stamos. View planning for site modeling. In *Proc. DARPA Image Understanding Workshop*, 1998.

[2] Elan Amir, Steven McCanne, and Randy Katz. Receiver-driven bandwidth adaptation for light-weight sessions. In *Proc. of ACM Multimedia*, Seattle, Washington, November 1997.

[3] R. Azuma. A survey of augmented reality. *Presence*, 6(4):355–385, August 1997.

[4] H. Beadle, B. Harper, G. Maguire Jr., and J. Judge. Location aware mobile computing. In *Proc. ICT '97 (IEEE/IEE Int. Conf. on Telecomm.)*, Melbourne, Australia, 1997.

[5] Vittoria de Nitto Persone, Vincenzo Grassi, and Antonio Morlupi. Modeling and evaluation of prefetching policies for context-aware information services. In *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM-98)*, pages 55–65, New York, October 25–30 1998. ACM Press.

[6] Mike Esler, Jeffrey Hightower, Tom Anderson, and Gaetano Borrielo. Next century challenges: Data-centric networking for invisible computing. In *Mobicom 99*, August 1999.

[7] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the world: 2D windows for 3D augmented reality. In *Proc. UIST '93 (ACM Symp. on User Interface Software and Technology)*, pages 145–155, Atlanta, GA, November 3–5 1993.

[8] S. Feiner, B. MacIntyre, T. Höllerer, and A. Webster. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. In *Proc. ISWC '97 (Int. Symp. on Wearable Computers)*, pages 74–81, Cambridge, MA, October 13–14 1997.

[9] Richard H. Frenkiel, B. R. Badrinath, Joan Borrs, and Roy D. Yates. The infostations challenge: balancing cost and ubiquity in delivering wireless data. *IEEE Personal Communications Magazine*, 7(2):66–71, April 2000.

[10] A. Gueorguiev, P. K. Allen, E. Gold, and P. Blaer. Design, architecture and control of a mobile site modeling robot. In *IEEE Int. Conf. on Robotics and Automation*, pages 3266–3271, April 24-28 2000.

[11] A. Harter, A. Hopper, P. Steggles, and P. Webster. The anatomy of a context-aware application. In ACM, editor, *Proc. of 5th ACM/IEEE Intl. Conference on Mobile Computingand Networking (MobiCom)*, August 1999.

[12] T. Höllerer, S. Feiner, and J. Pavlik. Situated documentaries: Embedding multimedia presentations in the real world. In *Proc. ISWC '99 (Third Int. Symp. on Wearable Computers)*, pages 79–86, San Francisco, CA, October 18–19 1999.

[13] T. Höllerer, S. Feiner, T. Terauchi, G. Rashid, and D. Hallaway. Exploring MARS: Developing indoor and outdoor user interfaces to a mobile augmented reality system. *Computers and Graphics*, 23(6):779–785, 1999.

[14] K. A. Hua and S. Sheu. Skyscraper boradcasting: A new boradcasting scheme for metropolitan video-on-demand systems. *ACM Computer Communication Review*, 27(4):89–100, October 1997. ACM SIGCOMM'97, Sept. 1997.

[15] Ana Lcia Iacono and Christopher Rose. Minimizing file delivery delay in an Infostations system. Technical Report WINLAB-TR-167, Rutgers University, Piscataway, New Jersey, August 1998.

[16] T. Imielinski and S. Goel. Dataspace - querying and monitoring deeply networked collections in physical space. In *Proc. of International Workshop on Data Engineering for Wireless and Mobile Access (MobiDE'99)*, Seattle, Washington, August 1999.

[17] H. Ishii, C. Wisneski, S. Brave, A. Dahley, M. Gorbet, B. Ullmer, and P. Yarin. ambientROOM: Integrating ambient media with architectural space (video). In *Summary of Conference on Human Factors in Computing Systems (CHI '98)*, pages 173–174, Los Angeles, CA, April 1998.

[18] Quinn Jacobson and Pei Cao. Potential and limits of Web prefetching between low-bandwidth clients and proxies. In *Proceedings of the Third International WWW Caching Workshop*, June 1998.

[19] B. MacIntyre and S. Feiner. Language-level support for exploratory programming of distributed virtual environments. In *Proc. UIST '96*, pages 83–94, Seattle, WA, November 6–8 1996.

[20] J. C. Navas and T. Imielinski. Geographic addressing and routing. In *Proc. of 3rd ACM/IEEE Intl. Conference on Mobile Computing and Networking (MobiCom)*, Budapest, Hungary, September 1997.

[21] Maria Papadopouli and Henning Schulzrinne. Seven degrees of separation in mobile ad hoc networks. In *globecom*, San Francisco, California, November 2000.

[22] Michael Reed and Peter K. Allen. 3-D modeling from range imagery. *Image and Vision Computing*, 17(1):99–111, February 1999.

[23] Michael Reed and Peter K. Allen. Constraint-based sensor planning for scene modeling. In *Computational Intelligence in Robotics and Automation (CIRA99)*, pages 131–136, November 7-9 1999.

[24] J. Rosenberg and H. Schulzrinne. A framework for telephony routing over IP. Request for Comments 2871, Internet Engineering Task Force, June 2000.

[25] Jonathan Rosenberg and Henning Schulzrinne. Internet telephony gateway location. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, San Francisco, California, March/April 1998.

[26] B. Schilit, N. Adams, and R. Want. Context-aware computing applications. In *IEEE Workshop on Mobile Computing Systems and Applications*, pages 85–90, Santa Cruz, CA, December 7-9 1994.

[27] Nachum Shacham. Multipoint communication by hierarchically encoded data. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, volume 3, pages 2107–2114 (9A.4), Florence, Italy, May 1992. IEEE.

[28] A. Smailagic and D. Siewiorek. The CMU mobile computers: A new generation of computer systems. In *Proc. COMPCON '94*, pages 467–473, February 1994.

[29] M. Spreitzer and M. Theimer. Providing location information in a ubiquitous computing environment. In *14th ACM Symposium on Operating System Principles*, volume 27, December 1993.

[30] Ioannis Stamos and Peter K. Allen. 3-d model construction using range and image data. In *Computer Vision and Pattern Recognition Conference (CVPR)*, pages 531–536, June 13-15 2000.

[31] Ioannis Stamos and Peter K. Allen. Integration of range and image sensing for photorealistic 3d modeling. In *IEEE Int. Conf. on Robotics and Automation*, pages 1435–1440, April 24-28 2000.

[32] T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healey, D. Kirsch, R. Picard, and A. Pentland. Augmented reality through wearable computing. *Presence*, 6(4):386–398, August 1997.