Ubiquitous Annotation Systems: Technologies and Challenges

Frank Allan Hansen
Center for Interactive Spaces,
Department of Computer Science, University of Aarhus
Åbogade 34, DK-8200 Århus N, Denmark
fah@daimi.au.dk

ABSTRACT

Ubiquitous annotation systems allow users to annotate physical places, objects, and persons with digital information. Especially in the field of location based information systems much work has been done to implement adaptive and context-aware systems, but few efforts have focused on the general requirements for linking information to objects in both physical and digital space. This paper surveys annotation techniques from open hypermedia systems, Web based annotation systems, and mobile and augmented reality systems to illustrate different approaches to four central challenges ubiquitous annotation systems have to deal with: anchoring, structuring, presentation, and authoring. Through a number of examples each challenge is discussed and HyCon, a context-aware hypermedia framework developed at the University of Aarhus, Denmark, is used to illustrate an integrated approach to ubiquitous annotations. Finally, a taxonomy of annotation systems is presented. The taxonomy can be used both to categorize system based on the way they present annotations and to choose the right technology for interfacing with annotations when implementing new systems.

Categories and Subject Descriptors: H.5.4 [Information Interface and Presentation]: Hypermedia.

General Terms: Design, Experimentation, Documentation, Standardization, Human Factors.

Keywords: Annotation, Ubiquitous Hypermedia, Mobile Computing, Context-aware Computing.

1. INTRODUCTION

For centuries annotations have been used as a tool by readers to personalize and contextualize written texts. Underlined or highlighted words are ways to summarize texts at a glance while marginalia and sketches have been used as reminders, personal discussions, and as communication media when texts were given from one person to another.

When text became digital, much emphasis was put on supporting the annotative activities in the new medium: as early as 1945 Vannevar Bush proposed the Memex [8] which would let readers comment on existing material and contextualize the material by linking information items into associative trails; and Douglas Engelbart implemented commentary facilities on the journal system in Augment [17] in the 1960s.

Much effort has been put into developing digital annotation facilities, especially in the hypermedia and Web communities. Open hypermedia systems such as Microcosm [13], Devise Hypermedia [20], and Chimera [2] support documents to be augmented with externally stored hypermedia structures (i.e. annotations and links) allowing users to create both personal and shared layers of information on top of the document corpus. Other systems like DLS [9], Webvise [21], and Arakne [3] focus on close integration with the Web. Pure Web based annotation tools include Xspect [12], the Annotea system from W3C [33], ComMentor [45], and the iMarkup plug-in [32].

However, not only text has been annotated. Since the early cave paintings people have enjoyed marking physical objects with information. Even today, graffiti painters tag public spaces with their signature to state their presence. In a less obtrusive way 3M Post-it notes allow physical objects to be augmented with information.

Just as the computer added machine support for annotations and links to the written text, the emergence of mobile and ubiquitous computer systems has spurred the development of computer mediated physical annotation systems. Common to many of these systems is the support for users to access and create information related to their context (i.e. their location in physical space, the objects they are working with, the task at hand, and so forth). Instead of placing information directly on the physical objects, these systems rely on information stored on remote servers, and attached to the physical objects by means of various context sensors.

HT'06, September 6–9, 2005, Odense, Denmark.

Different approaches have been taken to implement ubiquitous annotation systems; Cyberguide [1], Stick-e notes [39], Geonotes [18], comMotion [37], HP Websigns [40], and HyConExplorer [29][4] have been implemented on wireless, handheld devices like PDAs and smart phones while for example M.A.R.S. [31] and Sony's NaviCam [42] use wearable computers and augmented reality techniques to annotate and augment places and objects. HyConExplorer and the Digital Graffiti system [10] are examples of systems taking a hybrid approach where handheld devices are used to create the annotations in situ while the information can be presented on large public displays and other room components.

1.1 Four Challenges for Ubiquitous Annotations Systems

Despite these efforts little research exists which describe the general requirements for ubiquitous annotation systems and the challenges involved in implementing such systems. If information cannot be attached physically to an object how can physical entities (persons, places, and objects) be annotated and augmented with digital information? What technologies are required? What input and output devices are suitable for creating and presenting annotations in the physical world?

By investigating the lessons learned from open hypermedia and Web based annotation systems and the more recent work from augmented reality, wearable and ubiquitous computing this paper focuses on the requirements, problems and prospects for ubiquitous annotations. Different approaches are described through a number of system examples and HyCon, a context-aware hypermedia framework developed at the University of Aarhus, is used to illustrate an integrated approach to ubiquitous annotations. The investigation is centered on four important challenges:

Anchoring: The basic requirement for linking information and resources is that the resources can be *identified*. Different technologies can be employed for identification, but the precision of the technology determines how well annotations can be placed in relation to the annotated resources.

Structure: As many objects cannot be annotated directly, the structure of annotation objects (the object relationships) has to be general enough to allow any object, that have been identified to be annotated and linked. The structure also has to be general enough so that different anchoring techniques (and combinations thereof) can be employed.

Presentation: Information can be presented in close relation to the annotated resource e.g. as marginalia in a written text or graphics superimposed on physical objects. It may also be detached from the annotated object as text on a PDA, or a smart phone informing the user of the current location (i.e. the annotation is not presented on the location but in conjunction with it). Finally, information can also be completely detached, if presented "off location", separately from the annotated resource, e.g. on a remote PC that displays the annotation together with references to the annotated objects such as maps, photos, or descriptions.

Editing: It may not always be convenient to use the presentation interface for editing and authoring of annotations and more often than not, multiple interfaces are required e.g., when authoring material on a mobile device and presenting it in a Web browser on a PC or visa versa. Therefore, the architecture and infrastructure of the annotation system must be extendable and support integration of multiple systems and devices.

The paper is structured as follows: Section 2 discuss how annotations can be anchored in both digital and physical resources, Section 3 focuses on the link structures needed to link the annotations to real world objects, and Section 4 and 5 discuss methods to present annotations and the link relationships. The discussions in these sections take outset in a taxonomy for presentation and authoring in the field. Finally, Section 6 discusses the design of HyCon applications in relation to the taxonomy and Section 7 concludes the paper.

2. ANCHORING

Analog annotation systems typically allow information to be positioned directly on or within the annotated resource. In a book, words can be highlighted and the annotation can be placed between the lines or in the margin. A physical Post-it note can be placed directly on the objects it comments on and its small size even allow multiple notes to be placed side by side to give detailed information on smaller identifiable parts of a larger object e.g. on different areas of a whiteboard.

Precise identification and anchoring have several advantages: When an annotation is presented together with the annotated object there is no need to describe the object in the annotation itself thus reducing the effort of the annotator when creating the annotation. This is both true for written annotations in text, e.g. an underlined word in the text may serve as an anchor for an annotation in the margin, and then there is no need to rewrite the annotated word or its context in the comment, but is also true for physical objects. A Post-it note with the text "read chapter 4 and 5 in this" attached to a book does not have to reference the author of the book or the book's title to explain, which chapters the annotation refers to. This is also called a *spatial deixis* [18], since the interpretation of the annotation depends on the spatial context of the Post-it note, which in this case is on the books cover. In semiotic terms, the strong relation between the information in the annotation

Table 1. Examples of methods for identifying physical anchor points for annotations.

	Method	Hardware	Accuracy	Technical feasibility	User effort	Availability	Examples of use
Positioning	Manual input	Present	Mixed[a]	Easy	High	Every where	[19][50][52]
	IR sensors	Present	Room based	Easy	Low/medium	Indoors	[51]
	Ultrasonic positioning	Emerging	Room based (ID[c])	Medium	Low	Indoors	[30][41]
	WLAN positioning	Emerging	Room based (ID[c])	Medium	Low	Indoors[b]	[11][18]
	GSM positioning	Emerging	1-10km (ID[c])	Medium	Low	Every where	[49]
	Bluetooth positioning	Emerging	1-10m (ID[c])	Medium	Low	Indoors[b]	[6]
	GPS	Present	10m-50m	Easy	Low/medium	Outdoor	[1][29][37][40]
	Differential GPS	Present	1cm-1m	Medium	Low/medium	Outdoor	[30]
Tagging/ID	Barcodes	Present	ID	Easy	Medium	Every where	[36]
	2D barcodes	Present	ID	Medium	Medium	Every where	[43][46][45][47][48]
	RFID	Present	ID	Easy	Medium	Every where	[24]
	IR Beacons	Present	ID	Easy	Medium	Indoors	[34][51]

a. System and user dependent

b. And outdoors in hot-spot areas

c. Device may be used as ID

and its spatial context is called an *indexical representation* as opposed to a symbolic or iconic representation, that is independent of time and place (e.g. a more formal reference to the chapters from another book) [35]. Indexicality is a property of the representation, that gives it a context-specific meaning, and therefore only makes sense in a particular setting. The more precise the annotation is anchored into a context, the easier it is for the annotator to use deictic expressions and write shorter messages relying on semiotic indexes.

However, when it comes to digital annotation systems, it cannot be assumed that the annotations can be added directly to the annotated resource: text cannot be added to documents on the Web, for instance, if the documents are write-protected and the annotator does not have write privileges to the document. Similarly, as most physical objects do not have computational capacities, it is impossible for them to contain digital annotations (to overcome this, some efforts like HP Labs' "Cooltown" [34] suggests adding Web servers to physical objects to be able to hold information about the objects and create a so called "Web presence" for the objects).

Nevertheless, as long as a resource (digital or physical) can be identified its identity can be used to specify the anchor point for annotations. This basic observation has been used as the foundation for anchor based hypermedia which relies on resources (content) and structures (links and annotations) to be separated but connected through anchor specification that identifies the exact part of a resource that is annotated.

2.1 Anchors for Digital Resources

Anchors have been formally defined in models like the Dexter Hypertext Reference Model [26] and the OHP-NAV model from the Open Hypermedia Systems Working Group [22]. Dexter introduced the notion of anchor objects containing anchor values while OHP-NAV used the term Location Specification (LocSpec). The actual format for the anchor values have been left opaque in the models since they should be general enough to specify any resource or part of a resource to be used as anchor point and this specification is of course highly dependent on the nature and structure of the resource. However, for a number of digital media formats, the anchor values specified by LocSpecs have been investigated in some detail [14][21][7]: fragments in a text document can be identified by a span in the text and for bitmap images a fragment can be identified by its pixel coordinates. In temporal media, such as sound or video, time-segments are useful fragment identifiers, and objects-IDs are natural candidates in object-oriented formats such documents produced by vector or CAD programs. Related work includes the XML Pointer Language (XPointer) from W3C [16] which supports several techniques for fine-grained anchoring in Web resources.

2.2 Anchors for Physical Resources

Techniques for identifying physical resources are also highly dependent on the nature of the resource. A physical location may be identified by its name, e.g. its postal address or building name, but can also be identified as a (x, y, z)-point in space given its latitude, longitude, and height above sea level. A person wearing an active badge can be identified by the badge ID, or simply by his or her name, a car can be identified by its license plate, and a book can be identified by its barcode. Examples of methods for acquiring these identifiers for physical resources are shown in Table 1.

Perhaps the simplest way of identifying physical resources (for the developers at least) is to let the user manually enter the location specification. But, as described in the table, this method requires most effort on behalf of the user. MS Research's WWMX [50] supports places to be annotated with photos but requires the user to specify the location by dragging the photo onto a map. Yellowarrow [52] and Grafepedia [19] include special physical markers (yellow arrows and blue underlined text) that contain a code the user has to enter and send in a SMS to receive the annotation.

However, identifying physical resources by using sensor equipment and tagging technologies provides a much better use experience, since it eases the task of both authoring and accessing the annotations.

2.2.1 Location Based Techniques

Location based information systems is probably the broadest group of physical annotation systems developed so far. One of the first location aware systems was prototyped on PARCTabs [51]. The PARCTab was developed at Xerox Parc in 1992 and took advantage of the simple idea of location awareness by using a palm-sized computer (the "tab") with a pen interface linked to an infrared network. The cellular network provided location information on a room-by-room basis. Several later systems have also been based on simple proximity techniques, e.g., the GUIDE [11] system calculates its position relative to 802.11 WLAN base stations to present location dependent information about Lancaster attractions. Geonotes [18] also uses WLAN positioning, but allow users to manually specify an extra *place label*—a user defined location specification that further specifies the anchor point of the annotation. Even though the place label has to be interpreted by the user it gives a much better description of the anchor point than the coarse grained measurement from the WLAN positioning system alone. On a larger scale, the GSM cellular net (and similar cell network technologies) can also be used to provide location information. Despite providing only coarse grain location information, cell information is used by some commercial, mobile annotation system e.g., TagandScan [49]. TagandScan users can annotate a location, determined from their current mobile phone cell, with text, image, and a category (e.g., restaurant), and publish this annotation to the general public, or just share it with their friends.

The Global Positioning System (GPS) is ideal for identifying outdoor locations, since it identifies places as global (x, y, z)-points. Cyberguide [1], HyConExplorer [29], ComMotion [37], HP Websigns [40], M.A.R.S. [31], and Topos [23] are systems using GPS coordinates as anchor points. The latter three combine the user's position with digital compass data to determine the position of anchor points relative to the user.

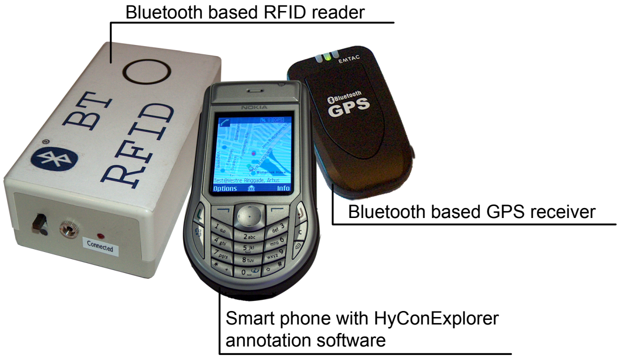


Figure 1: A mobile annotation system on a camera phone. Physical objects can be identified by their position (GPS) or ID (RFID or visual tags).

Table 1 lists GPS as being easy to implement: the infrastructure is readily available for anybody to use and GPS receivers have become a commodity. Often, some effort is required by the user to set up and connect the receiver. On the other hand, when using network based location technologies (Bluetooth, WLAN, cell-based, etc.) only little work has to be done by the user since the existing communication infrastructure is used (resulting in a lower user effort in the table). However, this infrastructure may not be available everywhere (e.g. Bluetooth or WLAN) or the location data may not be public available (e.g. from private cell networks). Furthermore, more sophisticated software is often needed to compute the location from measurements from multiple base-stations (through triangulations, scene analysis, etc.) and thus it may be harder to implement (resulting in a medium technical feasibility in the table). Another property that may influence the choice of anchor identification technology is the physical range of the technology. Radio based solutions cannot easily be confined within a fixed area such as a room but will penetrate walls and floors in buildings. On the other hand, infrared, ultrasonic, and the visual tagging technologies (discussed in the next section), cannot easily penetrate walls or furniture and is therefore limited to the physical area they are placed in. Which behavior is desirable depends on the annotation application; in some cases it may be useful only to see the anchors that are placed in the user's nearest context, e.g. the room the user is located in, and in other scenarios a broader view may be appropriate, e.g. if the user whishes to get an overview of comments on restaurants in the neighborhood.

2.2.2 Tagging Based Techniques

If object tagging can be employed, a number of other technologies can be used to identify the objects. Mackay's work on augmented paper [36] demonstrates how simple

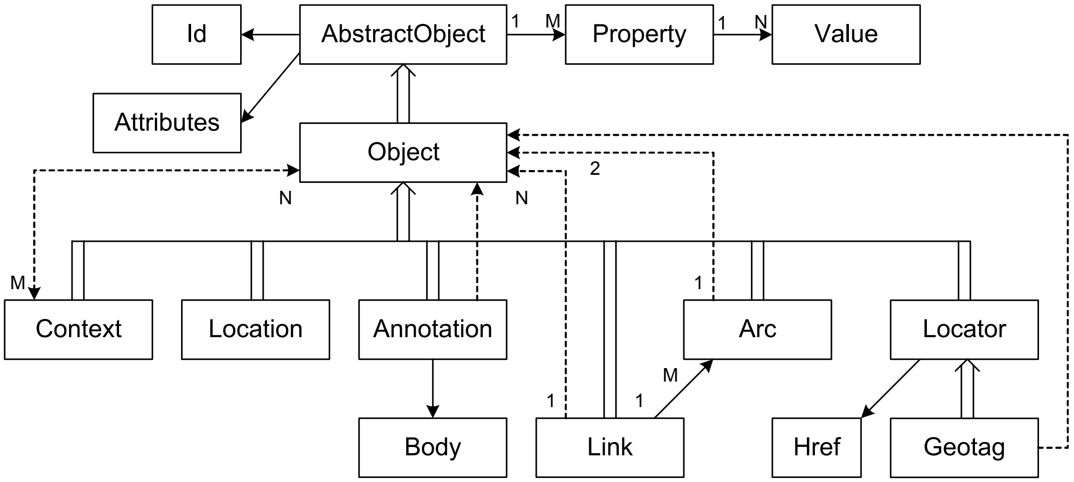


Figure 2: The HyCon data model (from [29]) encompasses hypermedia annotations, links, and context structures.

one-dimensional barcodes can be used to identify sheets of paper, which then can be annotated and augmented with digital information through a combination of cameras and projectors. More advanced visual markers include 2D-barcodes. Rekimoto's Cybercodes [43] can be recognized by a camera based reader, and used as the anchor point for augmented reality overlays. Semacodes [45] are also 2D barcodes, which support information to be encoded in the pattern. This is typically used to encode the URL to the linked information directly into the tag. Radio Frequency Identification (RFID) tags have the advantage over barcodes that they do not have to be visible. Small tags can be placed in books, on paper, or together with other physical material to identify the objects. Topos [24] is a physical hypermedia system designed to let landscape architects annotate physical objects (rocks, material samples, paper drawings, etc.) with digital information. RFID tags can be distributed with the objects, and the system utilizes the tags as anchor points for both information and links between the objects.

One example of a system that unifies a number of the technologies in Table 1 is the mobile annotation system HyConExplorer [29] illustrated in Figure 1. The system can use a combination of sensors to identify resources. Physical places can be identified by their latitude/longitude-coordinates from the GPS receiver and RFID tagged objects can be scanned and annotated with the RFID scanner. The application uses a Bluetooth link to communicate with the sensors, so other sensors can easily be added. The smart phone is equipped with a mega-pixel camera which potentially could be used to recognize visual tags, e.g. Semacodes [46] or High Energy Magic's Spotcodes [48].

3. STRUCTURE

While Section 2 discussed technologies for identifying objects, this section focuses on data models that utilize those anchoring technologies to tie the objects and the digital annotation information together.

Some systems rely on minimal structures, e.g. by annotating places with photos simply by using the support for GPS-coordinates in the EXIF header in JPEG-files (one may argue that this results in the photo being annotated with meta-information, however when used with an appropriate viewer this technique allows photos to be associated with locations). Other systems implement more advanced structures, but limit the annotation model to cover just one specific annotation medium, e.g. location based annotation systems that implements dedicated location models [18][37][50]. However, these models are very application specific and may be hard to apply to other domains, e.g. to applications relying on a range of the techniques discussed in Section 2. Thus, a more generic approach is generally desirable.

One such approach is the annotation model implemented in the context-aware Stick-e notes architecture [39]. Annotation objects in the Stick-e note model are defined by their content (the internal structure) and their context (the external structure). The context is recorded within the annotation to indicate the condition for which the annotation should be invoked. Different discrete contexts, depending on different trigger-conditions, can exist for a single annotation and cause the annotation to be presented, e.g. entering an office (location) or meeting a specific person (identity) or a combination thereof. This approach supports both physical and digital objects to be annotated as long as they can be identified and described in the models trigger-conditions.

However, one potential problem with this model is the combination of content and context in the annotation specification. As an example, consider the situation where a user wishes to annotate a location that has already been annotated by another user. If the system relies on location information from sensors, such as WLAN or GPS sensors, small inaccuracies in the measurements may result in the annotation being attached to a slightly different location than what the user expects (this was the reason why the reusable "place label" was introduced in Geonotes [18]). But reusing the existing annotation as the anchor point for the new annotation is also problematic: what is now being annotated? The location, or the existing annotation?

One solution to this problem is to separate the content and structure as described earlier and relying on an anchor based model. Open hypermedia models (e.g. OHP-NAV based models [22][38]) define an individual, identifiable object for each entity that is linked or annotated in the system (a node in the hypermedia network). The node will typically not contain any information itself, but simply be a proxy for the actual content. Different annotation objects can then be created and reference the same object. If the content has internal structure, further location specifications (LocSpecs) can be added for fine-grained anchoring.

Several hypermedia models have been generalized to support physical-digital links. The FOHM model [38] implements this by defining a single data item (a hypermedia node) that represents the entire physical universe. Aspects of the physical world are represented in this node and areas in the node can be referenced by name or location, just as fragments of text can be referenced in normal hypermedia documents. Physical and digital objects can in this way be interlinked by standard hypermedia mechanisms. Furthermore, the presentation of the annotations and links can be controlled by associating *context* objects with the hypermedia structures. These objects can be pattern matched against the user's current context to determine whether the annotation or link should be presented or not.

A similar approach is presented by the Dexter based HyperReal model [44]. In this model, the physical world is represented as specialized atomic components, e.g., physical space is presented by a *map component* with a built in coordinate system. The map component defines how links can be anchored in the component and support information to be linked to locations in the space the component specifies. Other aspects of the physical context are accessed through *Entity components* that model the connection between the physical and digital objects. External events can be sent to the HyperReal model which will trigger a *mode* (a context) that specifies what part of the structures (the entities) are accessible. This mechanism can be used to present links and annotations, when the user's physical context changes.

Our illustrating example, HyCon, implements an XLink-based [15] structure model (see Figure 2). Annotations on digital resources are modeled by Xlink Locator-objects which contain a URI to the annotated resource and Annotation-objects that reference the locator. The locspec-attribute of the Annotation-class can hold the anchor-values. Physical contexts and objects can be modeled in two ways: either as subclasses of the HyConObject-class (as Annotation-objects reference instances of the HyConObject-class) or directly as HyConObject instances that are parameterized with property-value(s) pairs (for a further discussion of these two approaches please refer to [28]). This generic support for describing both physical objects and context parameters makes it easy for developers to use the model to create annotation applications tailored to specific domains or to dynamically add new annotation media to the existing systems.

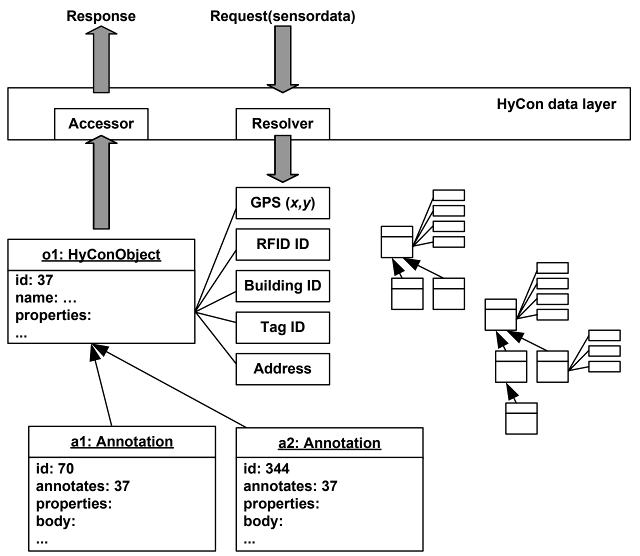


Figure 3: Resolving annotation structures from sensor data in the HyCon framework.

The use of an Xlink based model in HyCon also supports developers in choosing a pure Web based approach to expressing context information on the data objects. Xlink allows semantic meta information to be associated with the link elements (through the xlink:role attribute on link, resource, and locator elements and xlink:arcrole on arcs [15]). Using local, dynamic URIs, the semantic attributes can reference context elements that describe when and where the link element is visible and whether or not it is traversable in the user's current context. Thus, Xlink can be used as a lightweight and flexible data format for clients that understand the semantics of the link role, but even clients that are not capable of interpreting the context information (e.g. standard Web browsers) can still display the full link structures and let the user browse the information (independent of context).

3.1.1 Resolving Structures

Since annotations in this way can be structured according to their context, one could imagine that it also would be

Table 2: Taxonomy for ways to present annotations in relation to the annotated objects.

	Attached	**Detached**
On location	The user and object are co-located. Annotations are presented directly on the physical object. **Approach:** Augmented Reality (AR). **Technology:** VR helmets or goggles, built-in displays or speakers, projectors. **Examples:** [23][31][36][42]	Annotations are not presented on the annotated object but in conjunction with it. **Approach:** Ubiquitous Computing. **Technology:** Location or ID sensors. Mobile devices or public displays. **Examples:** [1][11][18][23][29][37][39][40]
Off location	Annotations are presented on a representation of the annotated object since the user and object are not co-located. **Approach:** Virtual Reality (VR). **Technology:** Computer models with information overlay. **Examples:** [47].	Annotations are presented only with a reference to the annotated object. **Approach:** Web presentation. **Technology:** Web pages. Online maps. References. **Examples:** [29][50][52]

possible to access the structures by giving a context specification.

In the Dexter Hypertext Reference Model, this kind of problem is handled by distinguishing between specifying and accessing objects in the model's storage layer [26]. These activities are realized through a *resolver* and *accessor* function. The resolver is responsible for locating objects (called components) by means of a component specification. This specification can either be a hard-wired component-ID or some constraint, which should be satisfied by the components (e.g., the component specifying a certain attribute/value pair). The result of the computation is a component-ID which is given to the accessor function which accesses and returns the actual objects.

Applying a similar abstraction to ubiquitous annotation models can be very powerful: data objects can be tagged with arbitrary context attributes that are used to anchor the data in a specific digital or physical context. When a user enters a physical context, sensor data measured in that context can be used to specify a request for relevant annotations. The request is given to the annotation system's resolver function and the resolver matches the context specification against the specification on the tagged data objects and returns the set of IDs of those objects that match the specification. Given the set of IDs, the accessor function can return the objects and their associated annotation and link structures to the client. The HyCon implementation of this mechanism is illustrated in Figure 3, but similar models exist in FOHM and HyperReal described in the last section.

4. PRESENTATION

Having presented how to structure annotations, we now turn to strategies for presenting and authoring the material.

Mackay [36] suggests three such strategies for augmenting the physical world with digital structures: 1) augment the user, 2) augment the physical object, and 3) augment the environment surrounding the user and object. The first approach requires the user to carry special I/O devices such as VR helmets and data gloves to see and interact with the digital information. The second approach relies either on electronics being built right into the physical object or on ubiquitous computing techniques using a combination of sensors to detect objects and mobile devices to present information associated with the objects (e.g. Xerox Parc's location system and tabs, pads, and boards devices [51]). The last strategy neither modifies the user or physical object, but lets the user interact with digital information through video projectors and cameras, that display information and capture the user's actions.

These three approaches could be adopted directly as the strategies for presenting annotation data in the world. However, Mackay's taxonomy is focused on augmented reality and assumes that the object of interest and the user are co-located and that the user interacts more or less directly with the physical object. This may not always be the case with annotation information, for example if a user has placed an annotation with recommendations on a new restaurant, she has just found, her friends may read this information "on location" on their mobile devices while passing by the restaurant but they could also be reading the information "off location" e.g. from a Web page, that integrates with the annotation system, and presents their friend's annotations.

So, instead of adopting Mackay's strategies directly we have chosen a slightly different taxonomy for presentation (depicted in Table 2). Annotations can be presented *attached* to the annotated object (i.e. presented on the annotation's anchor point) or it can be *detached* from the object. The first case can be accomplished by augmenting the user with a head-mounted see-through display and superimposing the annotation on the physical object but can also be done by augmenting the object directly with a presentation device (display, speaker, etc.). Augmenting the environment with projectors can also be used to present the annotations directly on the physical objects. Presenting information detached from the object can easily be done by utilizing a user's smart phone, PDA, or by using large stationary displays in the user's surroundings (corresponding to the notion of augmenting the user and augmenting the environment respectively). Besides these two modes, information can also be presented *on location* or *off location*. On location presentation requires the user to be co-located with the annotated object and corresponds to the examples just discussed, while off location presentation can occur when viewing the annotations on a Web page or displayed on a digital map on a PC. These four modes create a four-dimensional presentation space as illustrated in Table 2.

An interesting dimension is the one where annotations are presented off location, but attached to the object, since this might seem like a contradiction. However, virtual reality (VR) systems come close to this dimension as they do not present the actual physical object but a model of the object and annotation information can then be presented attached to the model. Table 2 lists the work from the University of Southampton [47] on contextualized open hypermedia as an example of such a VR model that has been augmented with annotations and controlled through a tangible interface. The table also cites examples for the other three approaches.

5. EDITING

Unfortunately, the majority of context-aware and location based information systems are "read-only". Information in these systems is pre-authored and users are reduced to readers without the ability to actively produce their own material. While this may be intended for some applications, e.g., navigation systems, tourist guides and the like [1][11], the goal of annotation systems is to let users both access and produce information. Today's mobile devices such as smart phones with built-in cameras, keyboards, and microphones, and with reasonable high network bandwidth capabilities are ideal hardware for running such systems. The phones can both be used to author and access annotations, and can thus serve as a powerful and creative tool for documentation.

However, it is not always easy to combine the presentation and authoring capabilities in the same application or with just a single technology. Even when it comes to mainstream and widely used technologies like Web browsers, users often have to resort to third party tools to be able to annotate and augment Web pages directly [3]. Some technologies are better suited for presentation than for authoring, e.g., the wearable computers used in M.A.R.S. [25] and for tangible hypermedia [47] are great for displaying 3D presentations on top of real world objects, but to create the presentation a separate desktop tool is needed. And sometimes, it is even desirable to split the tools used for presentation and authoring: a teacher preparing long and rich descriptions of places in the city, which her pupils should visit, will probably prefer to author the annotations using a desktop client rather than having to use a mobile phone. And conversely, the material created in the field by the pupils is probably better presented to the rest of the class on a Web page rather than on the phones used to author the annotations [5].

The four dimensions depicted in Table 2 can also be used to categorize different modes for annotation authoring, i.e., annotations can be created in situ either through direct manipulation of the objects or indirectly through some kind of device, or they may be authored off location and manually anchored in the digital or physical context. But as the discussion above illustrates an important requirement for the infrastructure is to support *integration* of the four different strategies in the taxonomy. Just like the architecture and data model in the systems should be general and extendable enough to encompass a number of different sensor types to allow annotations to be associated with different kinds of objects and anchored in different contexts, the architecture must also support a number of different mechanisms for both authoring and presentation.

6. EXAMPLE OF USE

The HyCon platform, which has been used as one of the illustrating examples in this paper, was designed with integration in mind from the beginning. In HyCon, services provide a common interface for resolving and accessing annotation and link structures. The services provide different interfaces and data formats and thereby allow clients with different capabilities to access the structures (for a more thorough discussion of the HyCon architecture please refer to [4][29]).

So far this approach has been very successful and a number of systems and services that utilize and share the annotation structures have been built. This includes mobile annotation systems for tablet PCs (HyConExplorer, depicted in Figure 4.a), smart phones (HyConExplorer/J2ME, 4.b), real time awareness applications for smart boards and interactive floors in school environments (HyConBoard, 4.c), and Web tools for both viewing annotations and for using the material to create online presentations and newspapers (HyConEditor, 4.d). Mostly using Web and mobile technologies, the HyCon applications can be placed in the right side of the taxonomy in Table 2, i.e. annotations are handled detached from the annotated resources. The tablet PC and phone applications support presentation and authoring "on location", juxtaposed to the physical objects being annotated, and the smart board and web interfaces supports "off location" presentation. The latter two differs with respect to temporal presentation: because the HyConBoard application is designed as an awareness component, other users must be using one of the mobile applications and actively be producing annotations for the board application to display the users' positions and present newly created annotations. The HyConEditor, on the other hand, functions as a purely virtual browser and editor, that does not require users or the author to be collocated with the annotated resources.

The taxonomy illustrates in this way that the current HyCon applications only utilize part of the design space for annotation interfaces. However, the HyCon annotation model is general enough to also encompass the left side of Table 2 as well. The research on systems like Topos [23], M.A.R.S. [31], and Sony's NaviCam [42] suggests that it may be desirable to have annotations presented directly attached or overlaid on the annotated resource.

One scenario that would benefit from this approach is systems that need to display information anchored on internal parts of physical objects that are not visible to the naked eye. This could be electrical wiring, plumbing in walls, bone structures in animals, and so forth. In a school setting for instance (which is the setting HyCon has been designed for), 2D visual tags (discussed in Section 2.2.2) could be placed on various objects and the children's camera phones could be used to track the tags and present information as overlays on the phones' displays. If the tags are placed on models in the classroom or figures in textbooks, the children would be able to investigate details of the models further through the information on the phones. This, in turn, would require systems that support the left side of the taxonomy, as well.

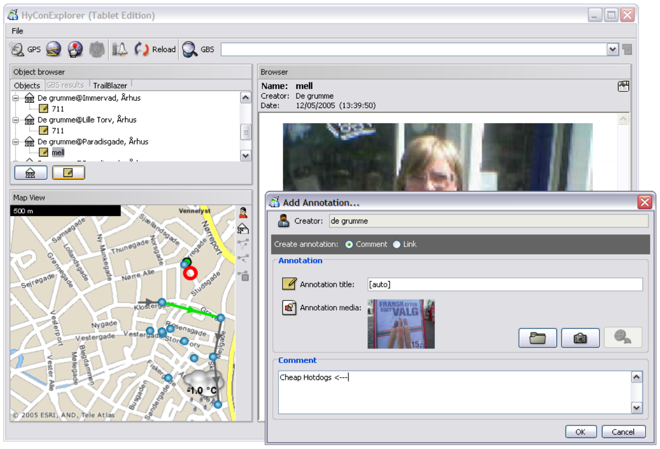
7. CONCLUSION

The emergence of mobile and pervasive computer technologies has brought computer-based annotations systems into the physical world. The combination of mobile and wearable computer devices, context-aware computing techniques, and open hypermedia and Web based annotation systems can be used to develop flexible and adaptive models for linking the digital and physical worlds. This paper has focused on four central challenges for implementers of ubiquitous annotation systems: anchoring, structuring, presentation, and authoring. Through a number of examples, it is demonstrated how sensed data can be used to both anchor, structure and resolve annotations in context. Finally, a taxonomy for ways to interface with the annotations is presented. The taxonomy can be used both to categorize system based on the way they present annotations and to choose the right technology for interfacing with annotations when implementing new systems.

8. REFERENCES

[1] Abowd, G. D., Atkeson, C. G., Hong, J., Long, S., Kooper, R., and Pinkerton, M. 1997. Cyberguide: a mobile context-aware tour guide. In *Wireless Networks* 3, 5 (Oct. 1997), 421-433. ACM.

[2] Anderson, K. M., Taylor, R. N., and Whitehead, E. J. 2000.

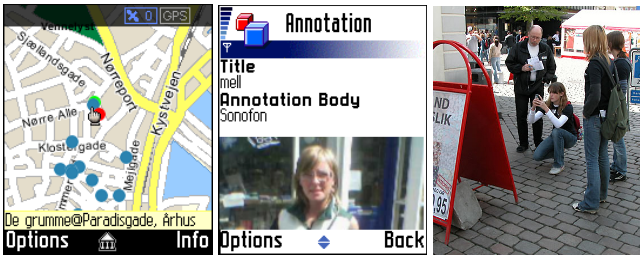


De grumme@Immervad, Århus
711
De grumme@Lille Torv, Århus
711
De grumme@Paradisgade, Århus
mell

Name: mell
Creator: De grumme
Date: 12/05/2005 (13:39:50)

Creator: de grumme
Create annotation: ● Comment ○ Link
Annotation title: [auto]
Annotation media:
Comment: Cheap Hotdogs <---

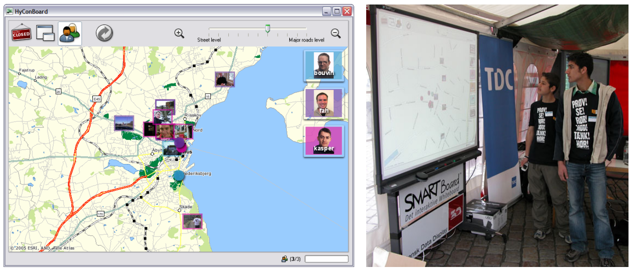
(a) Tablet PC interface



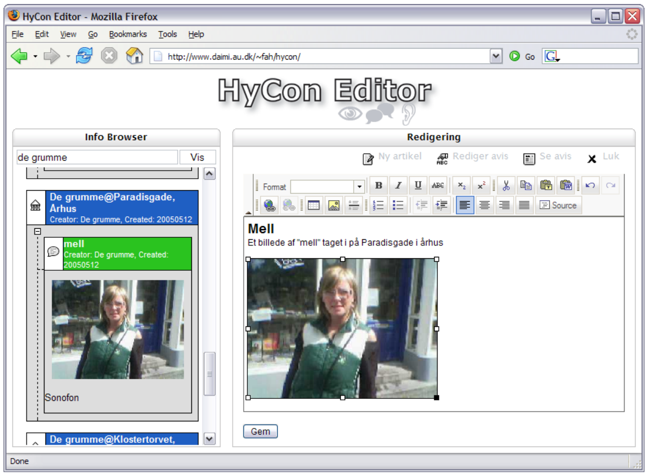
De grumme@Paradisgade, Århus

Annotation
Title
mell
Annotation Body
Sonofon

(b) Smart phone interface



(c) Smart board interface



http://www.daimi.au.dk/~fah/hycon/

de grumme

De grumme@Paradisgade, Århus
Creator: De grumme, Created: 20050512

mell
Creator: De grumme, Created: 20050512
Sonofon

De grumme@Klostertorvet,

Mell
Et billede af "mell" taget i på Paradisgade i århus

(d) Web authoring interface

Figure 4: Different interfaces for mobile authoring and presentation of annotations. All applications are integrated through common services (built on the HyCon framework, in this case) and shared structures.

Chimera: hypermedia for heterogeneous software development enviroments. *ACM Trans. Inf. Syst.* 18, 3 (Jul. 2000), 211-245.

[3] Bouvin, N. O. 1999. Unifying strategies for Web augmentation. In *Proc. of the 10th ACM Conf. on Hypertext and Hypermedia,* Darmstadt, Germany, February 1999. ACM Press, NY, 91-100.

[4] Bouvin, N. O., Christensen, B. G., Grønbæk, K., and Hansen, F. A. 2003. HyCon: A framework for context-aware mobile hypermedia. In *The New Review of Hypermedia and Multimedia (NRHM) journal volume 9*:59-88. Taylor & Francis, 2003, Abingdon, UK.

[5] Bouvin, N. O, Brodersen, C., Hansen, F. A., Sejer, O. I., and Nørregaard, P. 2005: Tools of Contextualization: Extending the Classroom to the Field**.** In *Proc. of the 4th Int. Conf. for Interaction Design and Children (IDC 2005)*, June 2005, Boulder, Colorado, USA.

[6] Brodersen, C., Christensen, B. G., Grønbæk, K., Dindler, C., and Sundararajah, B. 2005. eBag: a ubiquitous Web infrastructure for nomadic learning. In *Proc. of the 14th int. Conf. on World Wide Web* (Chiba, Japan, May 10 - 14, 2005). WWW '05. ACM Press, NY, 298-306.

[7] Brush, A. J., Bargeron, D., Gupta, A., and Cadiz, J. J. 2001. Robust annotation positioning in digital documents. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (Seattle, Washington, United States). CHI '01. ACM Press, NY, 285-292.

[8] Bush, V. As We May Think. *Atlantic Monthly*, Aug. 1945, pp. 101-108.

[9] Carr, L. A., DeRoure, D., Hall, W., and Hill, G. 1995. The distributed link service: A tool for publishers, authors and readers. *In Proc. of the 4th Int. World Wide Web Conf.*, Boston, USA, December 1995.

[10] Carter, S., Churchill, E., Denoue, L., Helfman, J., and Nelson, L. 2004. Digital graffiti: public annotation of multimedia content. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM Press, NY, 1207-1210.

[11] Cheverst, K., Mitchell, K., and Davies, N. 2002. The role of adaptive hypermedia in a context-aware tourist GUIDE. *Commun. ACM* 45, 5 (May. 2002), 47-51.

[12] Christensen, B. G., Hansen, F. A., and Bouvin, N. O. 2003. Xspect: bridging open hypermedia and XLink. In *Proc. of the 12th int. Conf. on World Wide Web*, Budapest, Hungary, May 20 - 24, 2003. ACM Press, NY, 490-499.

[13] Davis, H. C., Knight, S., and Hall, W. 1994. Light hypermedia link services: A study of third party integration. In *Proc. of the 1994 ACM European Conf. on Hypermedia technology*, Edinburgh, UK, pages 41–50.

[14] Davis, H. C. 1998. Referential integrity of links in open hypermedia systems. In *Proc. of the 9th ACM Conf. on Hypertext and Hypermedia.* Pittsburgh, Pennsylvania, United States, June 1998). ACM Press, NY, 207-216.

[15] DeRose, S., Maler, E., Orchard, D., and Trafford, B. (editors). XML Linking Language (XLink). W3C Rec. 27 June 2001, W3C, June 2001. http://www.w3.org/TR/xlink/.

[16] DeRose, S., Maler, E., and Daniel, R. 2001. XML Pointer Language (XPointer) Version 1.0. http://www.w3.org/TR/WD-xptr.

[17] Engelbart, D.C. 1984. Authorship Provisions in Augment. In *Proc. of 28th IEEE Int. Conf.*, San Francisco, CA, Feb. 27-Mar. 1, 1984.

[18] Espinoza, F., Persson, P., Sandin, A., Nyström, H., Cacciatore, E., and Bylund, M. 2001. GeoNotes: Social and Navigational Aspects of Location-Based Information Systems. In *Proc. of the 3rd int. Conf. on Ubiquitous Computing*, Atlanta, Georgia, USA, September 30 - October 02, 2001. Lecture Notes In Comp. Science, vol. 2201. Springer-Verlag, 2-17.

[19] Grafedia. 2005. Product page: http://www.grafedia.net/

[20] Grønbæk, K., Hem, J. A., Madsen, O. L., and Sloth, L. 1994. Cooperative hypermedia systems: a Dexter-based architecture. *Communications of the ACM* 37, 2 (Feb. 1994), 64-74.

[21] Grønbæk, K., Sloth, L., and Ørbæk, P. 1999. Webvise: browser and proxy support for open hypermedia structuring mechanisms on the World Wide Web. In *Proceeding of the 8th int. Conf. on World Wide Web*, Toronto, Canada, 1999. P. H. Enslow, Ed. Elsevier North-Holland, NY, 1331-1345.

[22] Grønbæk, K., Sloth, L., Bouvin, N.O. 2000. Open hypermedia as user controlled meta data for the Web. In *Computer Networks: The Int. Journal of Computer and Telecommunications Networking*, v.33 n.1-6, p.553-566, June 2000

[23] Grønbæk, K., Vestergaard, P. P., and Ørbæk, P. 2002. Towards geo-spatial hypermedia: Concepts and prototype implementation. In *Proc. of the 13th ACM Conf. on Hypertext and Hypermedia* (College Park, Maryland, USA, June 11 - 15, 2002). ACM Press, NY, 117-126.

[24] Grønbæk, K., Kristensen, J. F., Ørbæk, P., and Eriksen, M. A. 2003. "Physical hypermedia": organising collections of mixed physical and digital material. In *Proc. of the 14th ACM Conf. on Hypertext and Hypermedia* (Nottingham, UK, August 26 - 30, 2003). ACM Press, NY, 10-19.

[25] Güven, S., Feiner, S. 2003. Authoring 3D Hypermedia for Wearable Augmented and Virtual Reality, In *Proc. of ISWC '03 (7th Int. Symposium on Wearable Computers)*, White Plains, NY, October 21-23, 2003, 118-226.

[26] Halasz, F. and Schwartz, M. 1994. The Dexter hypertext reference model. *Commun. ACM* 37, 2 (Feb. 1994), 30-39.

[27] Hall, W., Davis, H., and Hutchings, G. 1996 *Rethinking Hypermedia: The Microcosm Approach*. Kluwer Academic Publishers.

[28] Hansen, F. A., 2005. Representing Context in Hypermedia Data Models. In *Proc. of the International workshop on Context in Mobile HCI in conjunction with MobileHCI'05*. Salzburg, Austria, September 19th, 2005.

[29] Hansen, F. A., Bouvin, N. O., Christensen, B. G., Grønbæk, K., Pedersen, T. B., and Gagach, J. 2004. Integrating the web and the world: contextual trails on the move. In *Proc. of the 15th ACM Conf. on Hypertext and Hypermedia*, Santa Cruz, CA, USA, August 09 - 13, 2004. HYPERTEXT '04. ACM Press, NY, 98-107.

[30] Harter, A., Hopper, A., Steggles, P., Ward, A., Webster, P. 1999. The Anatomy of a Context-Aware Application. In *Pro-*

c. of the 5th Annual ACM/IEEE Int. Conf. on Mobile Computing and Networking, MOBICOM'99, Seattle, Washington, USA, August 1999, pp. 59-68.

[31] Höllerer, T., Feiner, S., and Pavlik, J. 1999. Situated documentaries: Embedding multimedia presentations in the real world. In *Proc. of ISWC '99 (Int. Symposium on Wearable Computers)*, San Francisco, CA, October 18--19 1999.

[32] iMarkup: Annotate, organize and collaborate on the Web. http://www.imarkup.com/products/annotate_page.asp

[33] Kahan, J. and Koivunen, M. 2001. Annotea: an open RDF infrastructure for shared Web annotations. In *Proc. of the 10th int. Conf. on World Wide Web* (Hong Kong, Hong Kong, May 01 - 05, 2001). WWW '01. ACM Press, NY, 623-632.

[34] Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., and Spasojevic, M. 2002. People, places, things: web presence for the real world. *Mob. Netw. Appl.* 7, 5 (Oct. 2002), 365-376.

[35] Kjeldskov, J. 2002. “Just-In-Place” Information for Mobile Device Interfaces. In *Proc. of MobileHCI 2002*, Pisa, Italy. Lecture Notes in Computer Science, Berlin, Springer-Verlag.

[36] Mackay, W. E. 1998. Augmented reality: linking real and virtual worlds: a new paradigm for interacting with computers. In *Proc. of the Working Conf. on Advanced Visual interfaces* (L'Aquila, Italy, May 24 - 27, 1998). AVI '98. ACM Press, NY, 13-21.

[37] Marmasse, N. and Schmandt, C. 2000. Location-Aware Information Delivery with ComMotion. In *Proc. of the 2nd int. Symposium on Handheld and Ubiquitous Computing*, Bristol, UK, September 2000. Lecture Notes In Computer Science, vol. 1927. Springer-Verlag, London, 157-171.

[38] Millard, D. E., Michaelides, D. T., De Roure, D., and Weal, M. W. 2002. Beyond the Traditional Domains of Hypermedia. In Proc. of the 2002Workshop on Open Hypermedia Systems, ACM Hypertext ’02, Maryland, U.S.A.,June 11-15. Tech. Report from FernUniversitaet Hagen,Germany, 2002.

[39] Pascoe, J. 1997. The stick-e note architecture: extending the interface beyond the user. In *Proc. of the 2nd int. Conf. on intelligent User interfaces*, Orlando, Florida, United States, January 1997. IUI '97. ACM Press, New York, NY, 261-264.

[40] Pradhan, S., Brignone, C., Cui, J.-H., McReynolds, A., and Smith, M.T. 2001. Websigns: hyperlinking physical locations to the web. IEEE Computer, 34(8):42--48, 2001.

[41] Priyantha, N. B., Chakraborty, A., and Balakrishnan, H. The Cricket location-support system. In *Proc. of the Sixth Annual ACM Int. Conf. on Mobile Computing and Networking*, Boston, MA, August 2000. ACM Press.

[42] Rekimoto, J. and Katashi N. 1995. The World through the Computer: Computer Augmented Interaction with Real World Environments. In *Proc. of UIST '95 (8th Annual Symposium on User Interface Software and Technology)*, November 14-17, 1995, pp. 29-38.

[43] Rekimoto, J. and Saitoh, M. 1999. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems: the CHI Is the Limit* (Pittsburgh, Pennsylvania, United States, May 15 - 20, 1999). CHI '99. ACM Press, New York, NY, 378-385.

[44] Romero, L. and Correia, N. 2003. HyperReal: a hypermedia model for mixed reality. In *Proc. of the 14th ACM Conf. on Hypertext and Hypermedia* (Nottingham, UK, August 26 - 30, 2003). HYPERTEXT '03. ACM Press, NY, 2-9.

[45] Röscheisen, M., Mogensen, C., and Winograd, T. 1995. Beyond browsing: shared comments, SOAPs, trails, and on-line communities. In *Proc. of the 3rd int. World-Wide Web Conf. on Technology, Tools and Applications* (Darmstadt, Germany). P. H. Enslow and D. Kroemker, Eds. Elsevier North-Holland, New York, NY, 739-749.

[46] Semacode. 2005. Product Page: http://semacode.org/

[47] Sinclair, P., Martinez, K., Millard, D. E., and Weal, M. J. 2002. Links in the palm of your hand: tangible hypermedia using augmented reality. In *Proc. of the 13th ACM Conf. on Hypertext and Hypermedia* (College Park, Maryland, USA, June 2002). HYPERTEXT '02. ACM Press, NY, 127-136.

[48] Spotcode. 2005. Product page: http://www.highenergymagic.com/spotcode

[49] TagAndScan. 2005. Product Page: http://www.tagandscan.com/

[50] Toyama, K., Logan, R., and Roseway, A. 2003. Geographic location tags on digital images. In *Proc. of the Eleventh ACM int. Conf. on Multimedia* (Berkeley, CA, USA, November 02 - 08, 2003). MULTIMEDIA '03. ACM Press, NY, 156-166.

[51] Want, R., Schilit, B. N., Adams, N. I., Gold, R., Petersen, K., Goldberg, D., Ellis, J. R., and Weiser, M. 1995. An overview of the parctab ubiquitous computing environment. *IEEE Personal Communications*, 2(6):28--43, 1995.

[52] Yellow Arrow. 2005. Product page: http://yellowarrow.net