AUGMENTING HUMANS' EXPERIENCES

Luc JULIA

CHIC!, SRI International, Menlo Park. USA.
email: luc.julia@sri.com
http://www.chic.sri.com

ABSTRACT

In the real world, some information is directly available (e.g., shapes, colors, sometimes names). But the reality does not deliver all the information about the current space (e.g., architect, dates for buildings, names for people or plants, abstract data), the current activity and its context. It is almost certain, though, with the huge growth of the Internet, that most of the needed information is available somewhere in the largest database with which we have ever experimented. In this paper, we present very natural ways to interact with this handful of information, which is available upon natural request or in a proactive way to anybody. We want to study these new phenomena of invisible and pervasive computing that will drastically augment humans' experiences. People with perceptive deficiencies may benefit more than anybody else from such a technology, but the prototypes we are building to illustrate the concept and conduct the studies are meant for the good of all humanity.

KEY WORDS

Augmented Reality, Information Space, Multimodal Interfaces, User Modeling.

1 INFORMATION SPACE

The surrounding space contains information that can be either local (context) or global (somewhere on the Internet, in our experience). The real space does not contain all the information that one can find about it. For instance, when we are in front of a landmark, we are not likely to find the architect's name and the date it was built. But, this information is probably somewhere in the Web. The virtual space is, in this case, a better resource than the real world. Virtual reality can also enhance our own reality. Projecting our real world coordinates in a virtual world can lead to a handful of information. Also, the user brings his or her model to the space (e.g, information available in the office, the house). By putting all those pieces together, we are building a gigantic information space that tells the story of invisible computing, where all the components are talking to each other toward the goal of fulfilling the user's needs.

2 HUMANS' AUGMENTATION AND MULTIMODALITY

Since all kinds of information are now available through the Web, it is interesting to study how to access them. We could consider several new methods and instruments, but our approach is more "natural". We do not want people to have to learn new languages or to be familiar with different codes or specific devices. We believe that the current communication channels humans use to interact (e.g., speech, handwriting, gestures) are enough to get and sort the needed information, and we do not want to ask for any additional learning process. In this intelligent space, it is then important for humans to be well understood. The fusion of the different channels that humans are using to communicate their will is a key component for good recognition. Multimodality is useful for both repairing uncompleted commands and for the richness of the dialog. For several years, we have been studying how those modalities are useful by themselves or in synergy [1].

The way the information is going to be delivered is interesting as well. We definitely don't want the user to be overwhelmed by inadequate or unnecessary data. Studying what, how and when to present information to the user must be part of the experiments we are putting together. Of course, the very interaction with the environment will change, because of its reactiveness and proactiveness, and it is by building such applications that we will be able to identify the changes. We believe that the augmentation comes from this incredible amount of easy-to-access data that is relevant to the current task or context, but we have to prove it.

3 BUILDING A TEST BED

To study how people are going to react to such an augmented world, we are putting together a test bed that consists of a set of working prototypes, each of them dealing with a specific instance of this new kind of interaction. These prototypes are built on components coming from our research and the commercial world. By using them, we identify the weaknesses at both the component and interaction levels. Feeding back the results of the evaluation allows a rapid improvement of each piece of the prototype and, consequently, of the prototype as a whole.

3.1 TravelMATE

TravelMATE is an instance of our MATE (Multimedia Augmented Tutoring Environment) project, which focuses on augmented reality (Figure 1).

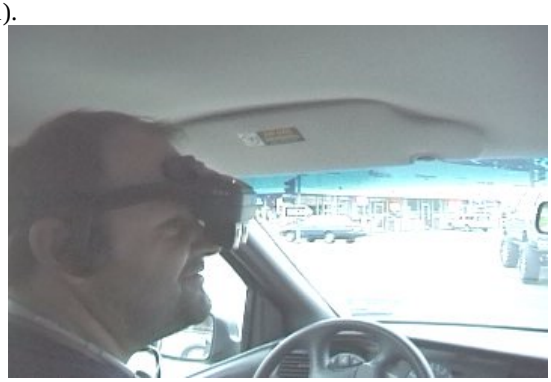

Figure 1: TravelMATE, augmented reality in cars

It is an extension of the CARS (Cooperative Agents and Recognition Systems) prototype we presented last year [2]. Driving a car requires a lot of attention, and the task of querying or accessing information should not add any cognitive load to the user. Using natural ways to access information is the goal of our study. When the driver wants

to know the location of interesting landmarks around his current position (given by GPS), he or she just asks for it: "Where are the interesting sites?". The system displays, with a transparent, very light background, a floating label at the location of each site. The size of a label gives an idea of the distance to the landmark. Then the driver, thanks to a simple gesture (3D gesture recognition) toward the label on the windshield, can select a specific site. Symmetrically, the windshield is proactive and, according to the user's interests or to some events happening in his or her space, will display accurate information or will deliver audio messages.

One of the interesting problems we are studying, thanks to this prototype, is how to deliver information just in time: the right information at the right moment, again without any additional cognitive load, in a very natural fashion for both the query and the delivery.

3.2 CHeF, the Collaborative Home eFridge

The CHeF, known also as the Smart Multimodal Fridge (Figure 2), focuses more on space awareness and interconnected appliances. From this Internet Fridge, one can order from a favorite online store (like Peapod or Webvan). Of course, the CHeF tracks all the items it contains as well as all the food items in the kitchen. It keeps up to date a local food database and a grocery list, and permanently verifies expiration dates for all the perishables. At the interface level, the multimedia display allows the user to write items on the grocery list (handwriting recognition) or to cross out (gesture recognition) items not wanted anymore. If anyone takes the last item of a kind, the CHeF speaks out and asks if it should add that to the grocery list. Since the CHeF is talking, it also accepts vocal commands (speech recognition) such as "what's for dinner tonight?" or "where is the rice?". The first query will search the local database and displays suggestions, and the latter will show on the screen a video of the kitchen, leading to the cupboard where the rice is stored.

Figure 2: CHeF, a Smart Multimodal Fridge

The CHeF is just a piece of the space in which the users' activities evolve. That's why it is connected to the rest of the world. In the current version it communicates with the car (TravelMATE) and a SmartTV (another piece of the test bed we didn't present here). We still want to add numerous abilities to the CHeF. It knows a lot about what is happening in and out of the kitchen, and it should know more about the user. We are planning to add some features (e.g., speaker identification, face recognition) that will personify the answers, but we want to stay away from "big-brother like" features.

4 FUTURE WORK

Putting together prototypes also allowed us to find out how close they are and that they reuse many components. We are actually in the process of tying those pieces even closer by building the CHIC! house, where it will be easy to observe people interacting with the prototypes in an even more realistic setting.

5 ACKNOWLEDGMENTS

Many thanks to our partners that make those projects possible (http://www.chic.sri.com/projects/partners.html) and to all the CHIC! people (http://www.chic.sri.com/people.html).

6 REFERENCES

[1] CHEYER, A. and JULIA, L. (1998). Multimodal Maps: An Agent-based Approach. *In Multimodal Human-Computer Communication, Lecture Notes in Artificial Intelligence #1374*, Springer, pp 111-121.
[2] JULIA, L. and CHEYER, A. (1998). Cooperative Agents and Recognition Systems (CARS) for Drivers and Passengers. OzCHI'98 : Adelaide (Australia), pp 32-38.