HABITATS²: A Cognitive Sensor System that Predicts Human Behavior

Matt Coler
INCAS³
Dr Nassaulaan 9
9401 HJ Assen
the Netherlands
MattColer@incas3.eu

CHI'12, May 5–10, 2012, Austin, Texas, USA.
ACM 978-1-4503-1016-1/12/05.

Abstract
HABITATS² is a network of cognitive sensors that will put a new real-time layer on our maps, taking us beyond the streets, houses, and rivers and provide us with the fluid essence of human experience: activity and behavior. Characterizing how people use an area based on a physical analysis of the acoustic and visual environment coupled with ethnographic and linguistic techniques to get at the nature of knowledge, experience and expectation drives the development of these cognitive sensors capable of linking the quantifiable external world and the internal complexities of the mind to maximize our environments, increase safety, and explore novel ways for humans and machines to interface. HABITATS² furthermore provides fertile grounds on which to launch further studies, bringing perceptual studies out of the lab and into the real world.

Author Keywords
cognitive sensors; sonic environment; affordances; symbiotic human-machine relations

ACM Classification Keywords
I.2.0 General: Cognitive Simulations; I.2.6 Learning: Knowledge Acquisition; I.2.9 Robotics: Sensors; J.4 Social and Behavior Sciences: Sociology

Introduction

HABITATS² (Human Activity in Built-environments and Intelligent Technology for Ambient Terrain Sensor System) is a network of cognitive sensors that will juxtapose a new real-time layer over existing maps, allowing us to remotely interact with and understand the environment in a way that goes beyond the limited view provided by static streets, houses, and rivers to provide a GIS visualization of the dynamic essence of human experience: activity and behavior. The resultant tool will provide real-estate developers, urban planners, city officials, and the police with the ability to remotely predict and monitor human activity in the northern Dutch city of Assen.

In the overwhelming majority of cases, efforts to characterize human use of and opinions about urban space use are achieved by administering closed questionnaires, and analyzing results alongside statistical and demographic information. The method used in developing HABITATS² represents a departure from such approaches. Instead, HABITATS² characterizes the shared use of space in terms of observed behavior and through open interviews targeted at identifying and characterizing the nature of the knowledge, experience and expectation relevant to behavior in a certain place.

Of key interest to the development of HABITATS², then, is the shared concept of public and private "space" in built environments and the use of that space for certain activities or sets of activities. Likewise, the nature of the interrelation between the physical and social realm [1] is of equal import. The exploration of these issues overlap with a range of social factors including, for example safety -- as manifest in e.g. Newman's work on defensible space [2].

In the simplest terms, HABITATS² focuses on identifying aspects of both the built and sonic environment relevant to human behavior alongside characterizations of relevant knowledge, experience, and expectation. Categories of physical and mental factors that interrelate to categories of activities for the basis for the model of an embedded AI program executed on each microphone.

Method

Summary description

Predicting something as complex as human activity requires a thorough description of the sonic environment and landscape (as mentioned in the previous section, the physical components of environments definitely influence behavior). This task is only possible with the pre-existing infrastructure provided by Sensor City in Assen. Sensor City made available a network of embedded microphones continuously recording the sonic environment throughout the city. This, once taken together with visual information relating to the landscape from satellite imagery and fly-over maps, supplies a good understanding of the external world which hosts the activities under investigation. The activities themselves will be documented using a range of ethnographic techniques *in situ* including participant observation and interviews.

In order to establish the correlation of sonic environment, built environment, and human activity, it is necessary to perform cognitive analyses. These analyses are targeted at linking categories of activity with physical components of the sonic and visual environment. The result of the cognitive analysis will be the establishment of rules (or, more aptly, "heuristics") that characterize the measured correlations between sonic environment, landscape, and activity. The definition of these rules will yield a by-product that is one of the cornerstones to this endeavor: the denotation of "sonic signatures" -- a particular blend of quantifiable sonic components that indicate that specific activities will likely be performed in a given space.

By referring to these rules and provided an appropriate characterization of the sonic signatures, we will develop embedded software that will transform the current Sensor City sensor network into an autonomous network of cognitive sensors. This will be done in a bootstrap process. The algorithms integrating all the information and data work on different quantities and qualities. As such, the algorithms need a basis both in artificial intelligence (be it, for example, expert systems, heuristics, or semantic networks) and machine learning. In this sense, HABITATS[2] represents a novel advancement of symbiotic human-machine relations insofar as it results in sensors that exhibit a level of spatio-temporal cognition in the real-world in a way similar to that of the human mind. Through the use of a GIS, it will be possible to integrate the Sensor City network, the embedded software, and the real-time data itself into an interface that provides users with the ability to remotely predict and monitor human activity in urban settings based only on environmental cues. Furthermore, from a more theoretical scientific perspective, this network will provide fertile grounds on which to launch further studies on, for example, the relationship between the acoustic signal and attention, the switching between hearing and listening modalities, and a host of other issues, effectively bringing perceptual studies out of the lab and into the real world.

Detailed overview of the research

This project is essentially a multidisciplinary endeavor designed to answer fundamental questions and ultimately model predictable aspects of human experience and cognition for use in the development of cognitive sensors. As such, it involves psychologists, linguists, ethnographers as well as physicists, programmers, AI researchers, with municipal workers, urban planners, and policy planners.

For clarity, the proposed endeavor can be broken down into five components. In the first, researchers perform a needs analysis of different end-users of the HABITATS[2]. As different end-users are interested in different aspects of human activity, they have different requirements. Accordingly, we will perform targeted discussions and semi-structured interviews with urban planners, city officials, the police, and policy advisors at the Municipality of Assen to determine which factors are of particular interest to each end-user group in terms of target locations, populations, physical parameters, and so on. The outcome of the interviews will also provide a preliminary idea of the appropriate "refresh rate" of the final interface as well as the grid-scale, physical parameters, and geographical locations of particular interest. The geographical locations will then be subjected to research to provide background in terms of the built environment, demographic information, and an overview of the general landscape. In the field,

Fieldwork performed by anthropologists and ethnographers **Concept**: Use ethnographic techniques to thoroughly document observed/expected human activity and experience of different users of the space at different points in time (see Fig. 1) and the physical environment. This qualitative exploration will be complemented with sound recordings and analysis to characterize previously-identified representative sites. **Techniques:** (a) participant observation; (b) open interviews, surveys; (c) focused descriptions of affordances in the environment that facilitate or inhibit documented behaviors; (d) documentation including photos and videos of human activity, preliminary recordings of sonic environment, etc.

fieldworkers will put particular emphasis on ethnographic techniques designed to thoroughly describe observed/expected human activity and experience of different users of the space at different times as well as qualities of the built environment and the affordances that encourage or inhibit behavior.

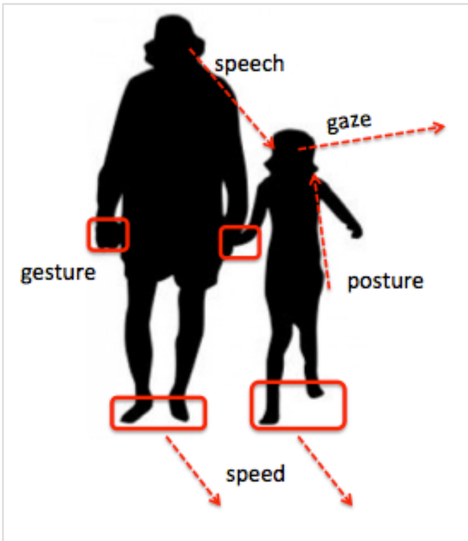

Figure 1: Some of the aspects of human behavior under analysis.

This qualitative exploration will be complemented with sound recordings and analysis to characterize representative sites previously identified. If relevant to the end-users, fieldworkers may engage in urban design analysis techniques like Kevin Lynch's map-making research methodologies. All such material gathered in this phase will be organized into digital dossiers which include field notes and observations about human activity and affordances, both of which are augmented by fly-over photographs provided by the municipality, relevant excerpts of transcribed interviews with users of the space, recordings, and satellite imagery.

Figure 2. Fieldwork reveals unexpected activity zones, as here a popular hangout spot on a sound-wall near a busy highway.

The next phase of the research is dedicated to finding correspondences between (1) human activities (for example, walking, exercising, reading, playing soccer), (2) outdoor spaces (for example, parks, plazas, intersections) and (3) the human perception of the environment (in terms of, for example, soundscapes, built environments, and affordances). The goal of this phase is to define a "sonic signature" (an auditory-based description characterizing a category of a sonic environment) for the categories of outdoor spaces identified by the end-user groups.

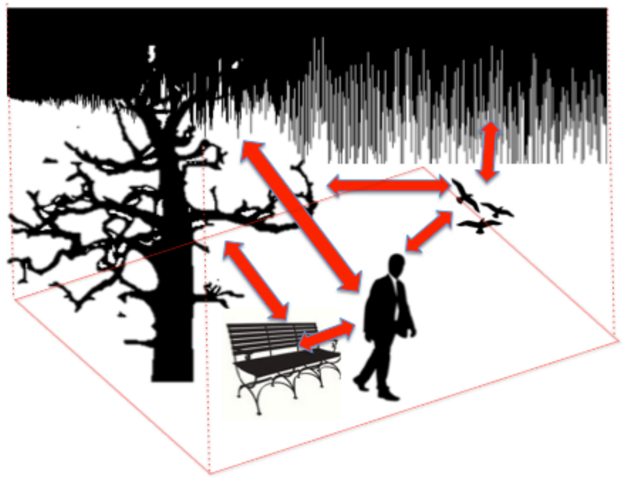

Figure 3. The definition of a given sonic signature is achieved by finding correspondences between human activities, outdoor spaces, and perception of the environment

Identifying the kinds of activities associated with the kinds of space can be achieved by linking activities (as performed by specific user-groups) and spaces as documented in the digital dossiers organized in the previous phase. Thereafter, a cognitive analysis on the types of activities and spaces will reveal categories of activities and spaces, thereby facilitating an analysis that will show associations between the two categories. Individuals from the user-groups will provide verbal descriptions of categories of space, activities, and perception of the environment via semi-structured or unstructured interviews, and oral descriptions of visual, audio and audiovisual representations of space. Linguistic analysis of this material will reveal the expectations of these subjects regarding what a particular space corresponding to a recorded sonic environment is optimally suited for. A study of the correspondences between the outcome of the linguistic analysis and the physical components in the environment (including streaming data from the Sensor City notes and visual elements as documented during fieldwork) will provide *a priori* kinds of sonic environments and built environments in relation to the activities and spaces defined. Next, the relevance of the *a priori* categories must be established. Thus, user-groups are subjected to cognitive tests, including free sorting tasks, triad tests, and prototype analysis, among others. Results will be analyzed with multidimensional scaling and phylogenetic tree analyses. At this point, it will be possible to define the physical components of the sonic and visual environment specific to each kind of space and activity. In the auditory modality, for example, the research will provide an "acoustic signature" (the characteristic components of the recorded sound in a given environment) for each kind. In other words, by this point, we will have established a link between mental representations with physical space and the sonic environment, thereby defining *a priori* kinds of soundscapes and built environments correlated to the activities and spaces previously defined.

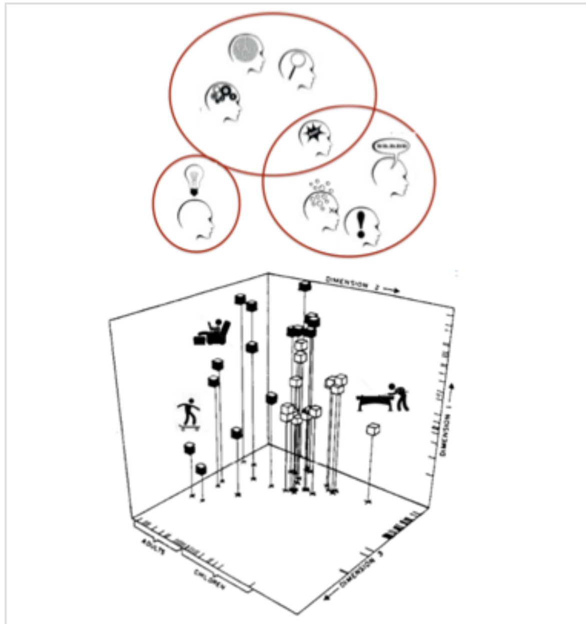

Figure 4. The output of cognitive tests like free sorting tasks (among others) can be analyzed with multidimensional scaling.

Following the establishment of this link, we will develop an acoustic signature that can be calculated automatically from audio recordings along with a classifier that can predict likely human activity based on the acoustic signature; that is, it will be capable of taking an acoustic signature and predicting a label that refers to some likely human activity. For example, a recording of a quiet park with the distant din of human chatter and the nearby singing of birds will yield a particular type of histogram (one with, among other things, high frequencies correlating to the birdsong) relevant components of will be linked to associated activities (e.g. reading or picnicking) in the classification step. To achieve this, it is first necessary to formalize the data gathered in the field into the following data types: activities, spaces, user-groups, and sonic-environments. Possible values (i.e. "instances of data types", so the data type "activities" would have, for example, values of "read", "play", or "shop") for those types arise from the cognitive tests. This information in hand, we will develop an algorithm that calculates the acoustic signature from an audio file. As alluded to previously, this may be achieved by making a cochleogram of the audio file, then making a histogram of all values in that cochleogram. This histogram is a generalized representation of the count content (e.g. loud values in high frequencies occur often, low-frequencies occur rarely). A different audio file would have a different histogram (e.g. one with low frequencies occurring more often). From this, it will be possible to develop a classifier capable of classifying likely activities for certain user-groups based on acoustic signatures. For example, we could use an acoustic signature (a histogram) to train a classifier to automatically assign a pre-established activity label to an audio file as a member of a certain user group would. Depending on the degree of success of the classifier, different algorithms for calculating the acoustic signature can be used. It is feasible to train the classifier to "learn" the association between the acoustic signature and the activity label, so that it can predict the label for a new sonic environment. Accordingly, a trained classifier could, for some unknown space, record the audio, calculate the acoustic signature, and in so doing predict human activity. Finally, the software for the acoustic signature calculator and classifier can be embedded on sensor nodes installed in the field.

In the penultimate phase, we gather data from the municipality, land registrar, bureau of statistics, and others. Using machine learning and other AI tools, we will build an automatic classifier based on the gathered data and categories. This refined classifier is capable of predicting human activity based on geographic data and the built environment (including affordances). It is then integrated with the previously-developed classifier (which predicts human activity based on acoustic signature). Thus, this combined classifier users not only GIS data, but also the output of the acoustic classifier to predict likely human activity.

Finally, in the last phase of the research, the goal is to implement a validated cognitive sensor network complete with a visualization methodology and deliver it to the end-users. The final product will include a real-time GIS visualization of the sonic environment and likely human activities with an interface tailored to the needs of the end-users. As an added benefit, this project will also yield a test-bed on which to launch studies on e.g. the acoustic signal and attention, the switching between hearing and listening modalities, differences in human perception and artificial information processing, etc. As such, from a theoretical vantage point, this research will lead to a better understanding of human perceptual identification and interpretation processes of sensor inputs or stimulations, and in so doing, bring us closer to extending more and more cognitivity to data processing in sensor systems. [3].

Demographic data is included in the visualization if relevant to the end-users.

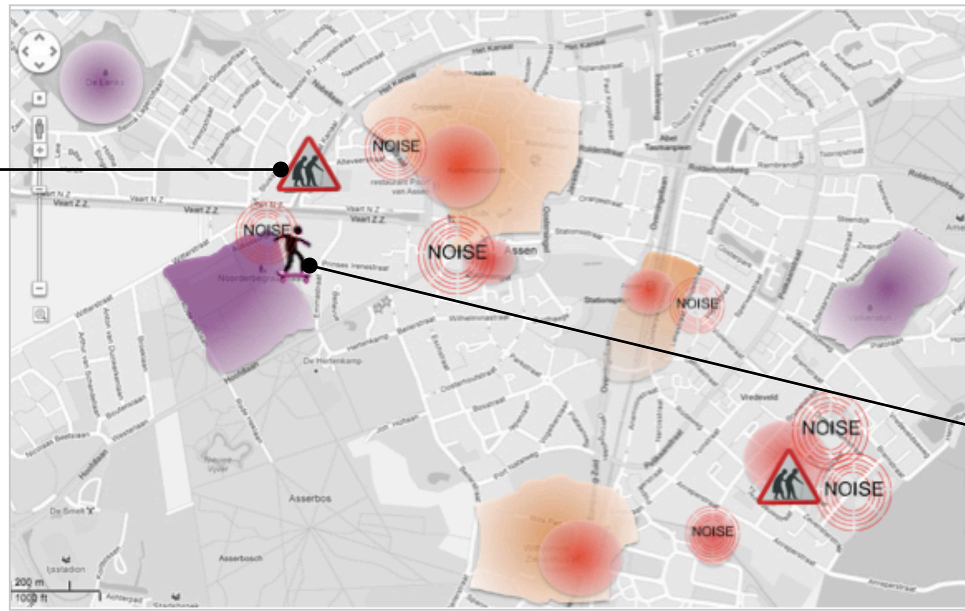

Figure 5. A hypothetical visualization of HABITATS[2] for a given point in time cross-referenced with noise complaints and locations with high density of senior citizens. The refresh rate depends on the needs of the end-user. Different color splotches represent different activities (see right sidebar).

The definitions of the activities arise from the output of the cognitive tests (including free sorting tasks, prototype analysis, and others).

Acknowledgements

This project was developed in close cooperation with the Cognitive Systems Team at INCAS³. Special thanks in particular to Dick Lyklema, Dirkjan Krijnders, and Gineke ten Holt. I am also grateful for the input of Julien Tardieu and Pascal Gaillard at PETRA (*Plateau d'Etudes Techniques et de Recherche en Audition*), CNRS, University Toulouse, Werner Kuhn (Institute for Geoinformatics. University of Münster), and Catherine Guastavino (McGill).

References

[1] Wilcox, B.L. & C.J. Holahan The social ecology of the megadorm in university student housing. *Journal of Educational Psychology* (1976).

[2] Newman, O. *Defensible space: crime prevention through urban design*. Macmillan, New York, 1972.

[3] Scientific Projection, INCAS[3], internal document, available on request (2012).