Design lifecycles and wearable computers for users with disabilities

Helen Petrie[1], Valerie Johnson[1], Stephen Furner[2], and Thomas Strothotte[3]

1. Sensory Disabilities Research Unit, Psychology Department, University of Hertfordshire,

Hatfield AL10 9AB, U.K.

2. Human Factors Unit, Centre for Human Communications Applied Research and Technology, British Telecom Laboratories, Martlesham Heath, Ipswich IP5 7RE, U.K.

3. Department of Simulation and Graphics, Otto von Guericke University of Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany.

Introduction

As with all technological artifacts, wearable and mobile computers need to be well-designed if they are to serve their functions appropriately. We are working towards an appropriate iterative user-centred design lifecycle for the development of wearable and mobile computers for people with visual disabilities, taking many ideas from mainstream HCI but adapting them both for the particular characteristics of wearable and mobile computer systems and the particular characteristics of our user group. This process has lead us to conclude that methodologies developed for the evaluation of static interfaces will need to be adapted and extended if they are to capture the critical features of peculiarities of wearable and mobile computers, whether they are for able-bodied or disabled users.

Wearable computers have enormous potential to assist people with disabilities. For example, for people with sensory disabilities such as blindness or deafness, they could provide substitute sensory information. Very cumbersome laboratory systems have been developed which provide substitute visual information for blind people by projecting a simple image in tactile form on the back or stomach, and these have been shown to have some utility [3]. Such systems would be far more useful as a wearable technology, although the appropriate miniaturization is still in the future. However, it is already possible to provide disabled people with useful information via wearable systems, even if this is not complete sensory substitution. For example, the Low Vision Enhancement System [13] is an augmented reality headset which helps the wearer make more effective use of any remaining vision by magnifying images and increasing light/dark contrast.

MoBIC: a wearable navigational aid for blind travellers

One example of our use of the iterative user-centred design lifecycle has been in the MoBIC Project, which has developed a wearable navigational aid for blind travellers [9, 12]. Blind people have two types of problems in moving through their environment, particularly if it is unfamiliar. Firstly, they need to avoid obstacles and to find a clear path to walk in their immediate environment (we have termed this micro-navigation [8]). This problem can be addressed remarkably well by a long cane or a guide dog, although travel may never be as easy as for sighted people. Secondly, blind people need to orient and navigate through the larger environment, which may require knowing which street they are on, which way they are facing, where to cross a street safely and so on. We have termed this macro-navigation. Without visual information, this macro-navigational problem can be enormously difficult even in familiar environments and it largely impossible in unfamiliar environments.

The MoBIC Outdoor System (MoODS), a wearable system, has been designed to assist in macro-navigation. It combines GPS and dGPS receivers with an on-board GIS which locates the traveller with reasonable accuracy on the digital map. The map contains not only standard information such as the street layout, house numbers and landmarks but also additional information of particular interest to blind travellers. As MoODS wearers move about, they need to interact with the system, being given appropriate information at appropriate times by a Trip Management System (TMS) to assist them in orientation and navigation. At times information needs to be user-initiated, for example when the user is uncertain of their location. At other times information needs to be system-initiated, for example to give warning.

The output from MoODS consists of synthetic speech messages which the user receives via headphones which are similar, but not identical, to Walkman-style headphones. The difference from standard Walkman-style headphones is that the MoODS headphones sit in front of the ears and do not cover them, thus they do not obscure auditory information coming from the environment which is vitally important to blind travellers. The input to the system is via a wrist-worn keypad with eight keys. The system can be worn in a backpack, in the pockets of a jacket or vest, or in a shoulder bag; the final prototype which was developed in the project weighted approximately five kilograms, with most of the weight being required for batteries. Figure 1 shows the shoulder bag version of MoODS. The input/output interfaces and their associated software were the end-product of an extensive sequence of user requirements elicitation studies and evaluations, which are outlined in the next sections.

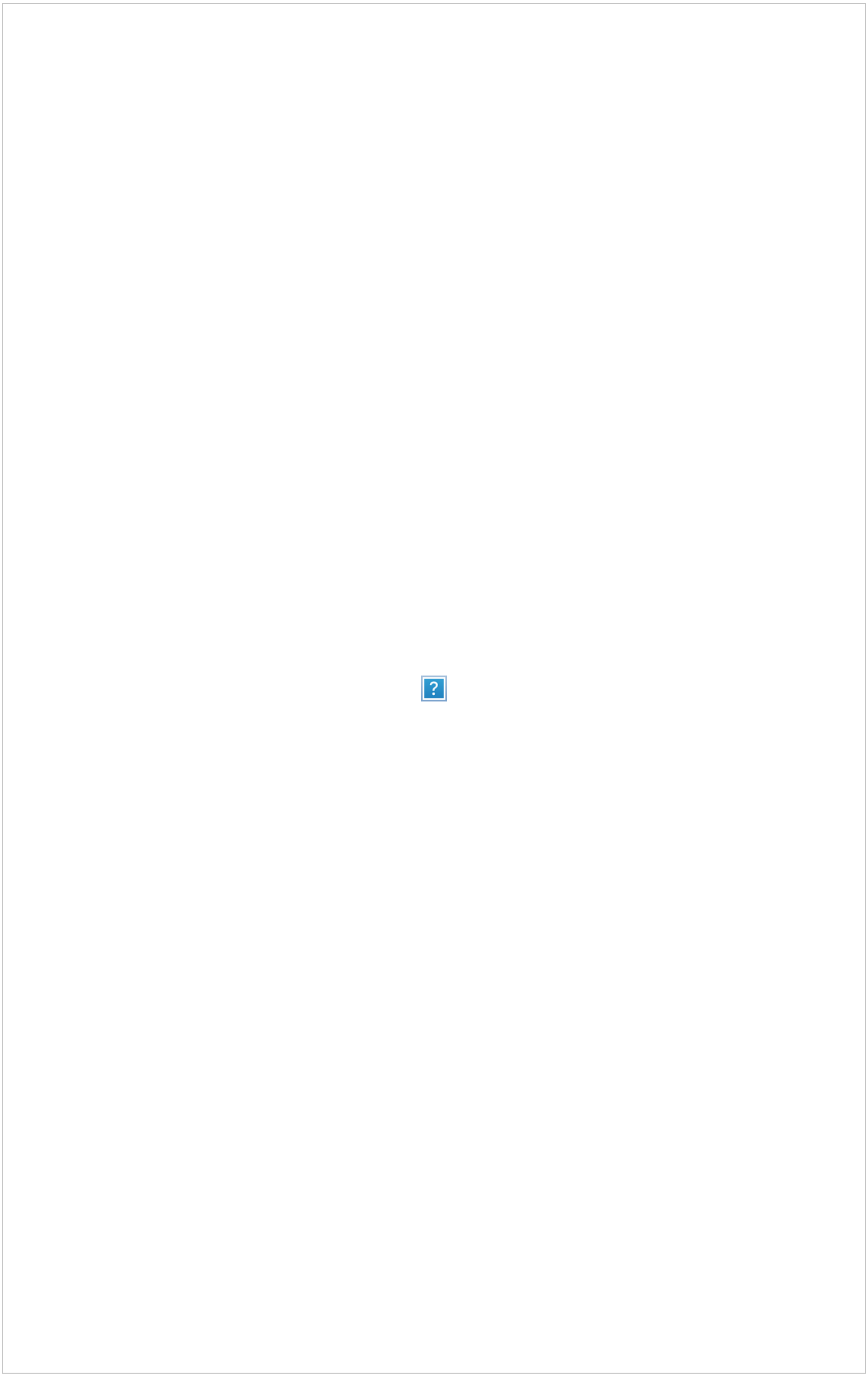

Figure 1: MoODS as it might be worn with a shoulder bag

User Requirements: from paper to cardboard and plastic prototypes

Classic methods of user requirements elicition were initially employed, with interviews and focus groups of potential users of the system and related professionals. However, it quickly became clear that everyone (including the design team) had enormous difficulty in imagining what using a MoODS might be like, both in terms of interaction devices and dialogue. This will be the case in any instance where a wearable is developed to perform a new function rather than simply undertake a known function in new, mobile contexts of use. However, potential users may also have difficulties imagining how they might undertake familiar tasks in new contexts.

In the MoBIC user requirements studies it was clear that participants and the design team were falling back on existing artifacts as metaphors for the use of the MoODS: it would be like a mobile telephone or a Walkman etc. While this can be helpful, it also limits the design space which is explored. In the case of MoODS, neither of these artifacts provided an adequate metaphor for the appropriate interaction. A mobile telephone style MoODS would be carried in a pocket and only interrogated when the user thought they required assistance. However, an important aspect of the functionality which the TMS can offer users is the provision of warnings. Contacting users via a phone call may be too slow to provide this information. A Walkman-style MoODS was a closer approximation to an appropriate metaphor of use, but blind travellers rely on auditory information from the environment, and wearing Walkman-style headphones would mask some sounds. In addition, users need to interact with the system more frequently than with a Walkman, so this metaphor provided no basis for the interaction with the system.

Exploring different metaphors of use and trying to invent new ones proved to be a useful method for potential users and the design team to clarify the MoODS design. A second successful method was the use of cardboard and plastic prototypes, the wearable answer to paper prototypes for 2D interfaces. In the case of input to MoODS, an initial prototype of a wrist worn keypad, similar to a watch, was presented to users along with several cardboard mock-ups representing a number of variations. These variations included different sizes for the keypad and different configurations and sizes of keys. Whilst users were not able to fully interact with these low fidelity prototypes they were able to judge what it would be like to wear them and how easy it would be to identify and operate the keys.

A third successful method for establishing the design was the use of simple "mobile Wizard of Oz" studies. For example, to establish the style for the basic navigational messages, a study was conducted with a short, typical inner city route (which involved turning corners, finding an appropriate point to cross a street and finding certain shops). The route was carefully studied and a suggested set of messages prepared. These messages were tape-recorded and potential users then walked the route with a sighted guide, the pre-recorded instructions guiding them from point to point along the route. At each point, the user paused and listened to the next instruction before acting on the message. Users were then asked to comment on the message structure, content and level of detail. In addition to providing information about how to formulate the navigational messages, this exercise also yielded useful information concerning physical interaction with MoODS and what users felt to be important in the design of input and output devices.

Our experience of user requirements elicitation for a wearable computer suggests that at least while wearable and mobile computers remain relatively rare, spending a lot of time and effort on classic user requirements elicitation techniques such as questionnaires, interviews and focus groups is not going to be particularly useful. We found that much more useful and interesting information could be obtained from potential users when they were given some kind of prototype or Wizard of Oz simulation of the system, even if this was only a very low fidelity version of the system or a component of the system. This was clearly because potential users found it too difficult to imagine something completely new and speculate about how they would like it to be - either physically or in its behaviour. However, when they were given simple concrete prototypes, these proved to be an extremely useful starting point for speculative discussions. In retrospect, we think the design team ought to have spent more time discussing the implicit metaphors of interaction (e.g. the mobile telephone and the Walkman) and having "metaphor busting" sessions, to explicitly break away from these existing metaphors and create new metaphors of interaction which might be more appropriate to our particular wearable system. Instead, we may have to wait until users master this particular wearable system and then find new tasks and situations in which to use it. Then we may well see new interaction styles begin to emerge, which as designers we ought to be analysing via a "psychology of tasks", as Carroll, Kellogg and Rosson [1] have proposed, to develop the next generation of designs for this wearable system.

Evaluation methodologies

In order to properly evaluate wearable and mobile computers, appropriate methodologies are needed. Exactly what tasks should be used and what measures are appropriate in such methodologies depends on the nature of the use of the wearable. Such use can be basically "serial" or "concurrent" in nature. In serial use, the use of the wearable alternates with other tasks, so that only one task is undertaken at any one moment. This is typical of applications such as the mobile desktop and online manuals: the user stops other activities to perform a task with the wearable. The more challenging and interesting situation is concurrent use, in which the user wishes to perform two tasks simultaneously, one of which involves the wearable. An increasingly common example of concurrent use of a wearable-like system is people driving cars while talking on a mobile phone (unfortunately, the latter are not usually designed as wearable devices which creates potentially dangerous situations).

In the case of serial use, evaluation needs to cover only the use of the wearable itself in appropriate contexts, and perhaps comparisons with the use of non-wearable equivalents. For example, for a wearable online manual, evaluation would cover the usability and acceptability of the system in situations in which use is proposed, and perhaps comparisons with more traditional situations of use such as consulting a manual in an office or library.

In the case of concurrent use, it is necessary to evaluate not only the use of the wearable itself, but also the other task which is being undertaken, to ensure that use of the wearable does not decrease performance on that task (this is the classic dual task paradigm of experimental psychology). In the MoBIC project the main focus of our evaluations has been on the usability and acceptability of MoODS. For this purpose a range of objective measures (e.g. system interrogations, errors) and subjective measures (e.g. rating scales of usability learnability, and satisfaction) have been developed. However, in addition, we have also investigated whether MoODS has any adverse effects on the user's concurrent task, that of micro-navigation. For this purpose users are asked to walk specially constructed routes and comparisons are made of their performance with and without MoODS.

Finally, it is also important to ensure that concurrent use of a wearable while undertaking another task does not put unacceptable stress or excessive workload on the users. In evaluations of MoODS we have taken objective and subjective measures of stress and used the NASA Task Load Index (TLX) [5] as a measure of workload.

Evaluation of MoODS

An evaluation of the final prototype of MoODS were conducted as part of the field trials of the MoBIC Project held in Birmingham in the summer of 1996. These took the form of intensive evaluation sessions in which controlled testing of the system and its acceptability took place. Seven blind participants took part, six men and one women, aged between 18 and 65 years. Six of the participants were congenitally blind and one was adventitiously blind. Participants were required to study two unfamiliar routes using the MoBIC Indoor System (a talking map and route planning system [9, 12]) and then when they felt sufficiently confident, to go out and walk these routes. In one condition this walk took place using both MoODS and the participant's primary aid (long cane or guide dog), in the other the walk was completed without MoODS and using only the primary aid.

A fully counterbalanced design was employed in which all participants walked one route with MoODS and one route without it. Each route was about one kilometre long and the two routes were matched as closely as possible for navigational difficulty. As MoODS is a wearable computer whose use is concurrent with another task, that of macro-navigation, and as that task is in itself a complex one, a wide range of measures was employed. Performance measures included time taken to walk the route, errors, number of system interrogations of MoODS and percentage preferred walking speed (PPWS). PPWS has been used as an evaluative measure for several electronic travel aids [2]. The rationale underlying the use of this measure is that all pedestrians have a walking speed which they prefer and that this appears to be the speed, which for them, is the most physically efficient. The ability of any mobility aid to allow the pedestrian to walk at this preferred walking speed (PWS) is therefore argued to be a measure of its effectiveness. It has also been suggested that since walking speed may be seen as an index of the degree of stress blind pedestrians experience, PPWS may also be used as a measure of stress [4,10].

Measuring psychological stress is not an easy task and finding a truly objective measure for it is especially difficult. Some studies have attempted this by measuring heart rate. For example Peake and Leonard [7] examined the heart rate of matched pairs of blind and sighted people when they walked guided and unguided. and they concluded that some form of psychological stress was responsible for the high heart rate of blind participants when walking unguided. However, heart rate is subject to change for a variety of reasons, so a subjective measure of stress was also used. The Spielberger State Trait Anxiety Inventory (STAI) [11] asks respondents to describe their stress/anxiety levels using 20 rating scales. It can be used to measure both trait (persistent, long term) stress/anxiety and state (transitory, short term) stress/anxiety, and both measures were used in this evaluation. The trait anxiety measure was used to investigate whether participants were generally anxious people, and the state anxiety measure to anxiety just before setting out on a walk and at a point midway during the walk.

When use of a wearable computer is concurrent with another task it is likely that cognitive workload will be increased and the TLX was included to measure this variable. The TLX is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales: Mental Demands, Physical Demands, Temporal Demands, Own Performance, Effort and Frustration. Three dimensions relate to the demands imposed on the subject (Mental, Physical and Temporal Demands) and the other three to the interaction of a subject with the task.

The perceived ease of use, efficiency, effectiveness, learnability and satifaction levels with MoODS were investigated via a number of 7 point rating scales and open-ended questions. Participants were also asked to rate how conspicuous they felt wearing MoODS, as this had been an issue raised during the user requirements studies. It is clearly important that a system is usable and that it is not difficult to learn how to use it, but it is equally important that users enjoy using it and are satisfied with it. On the surface, it would seem that if a system is easy to use, learnable, effective and efficient then any user will find it satisfying to use, but this is not always the case. For instance Nielsen. & Levy [6] note that there is often no correlation between subjective preference ratings and performance.

A summary of the measures used in the evaluation of the MoODS and the findings for these measures is given in Table 1, below. The only significant differences between navigation with and without MoODS were that the majority of participants achieved closer to their preferred walking speed with MoODS and that the perceived mental

Table 1: Summary of measures and findings from the evaluation of MoODS

Measure	Findings
Performance measures:	
Time taken to walk route	no significant difference with and without MoODS
Errors made during walk	no significant difference with and without MoODS
Percentage preferred walking speed achieved	4 out of 6 participants achieved closer to preferred walking speed with MoODS than without MoODS (data from final participant could not be analysed)
Stress-related measures:	
Heart rate	significantly higher during first third of walk, regardless of whether with or without MoODS
Trait anxiety	measured before setting out on one of the walks, relatively low mean scores
State anxiety	no significant difference with or without MoODS, either before or during the walks
Perceived cognitive load (TLX)	Mental demand significantly lower when walking with MoODS
Preference measures:	
Ease of use	Significantly higher than neutral rating
Learnability	Trend towards higher than neutral rating
Satisfaction	Neutral rating
Effectiveness	Trend towards higher than neutral rating

Efficiency	Neutral rating
Conspicuousness	Neutral rating

load of the task was lower when walking with MoODS. MoODS also received neutral or significantly positive ratings on the preference scales.

Conclusions

The practical constraints of conducting an evaluation which required training blind people in the use of a completely new technology meant that it was only possible to involve a relatively small number of participants. Thus it is difficult to draw any difinitive conclusions about the evaluation methodology itself. However, all of the measures contributed to our understanding of the use of MoODS. We believe that a number of new measures have been introduced in the evaluation which have potential in the evaluation of many wearable systems.

Appropriate iterative user-centred design lifecycles for wearable technologies need to be developed. These can build on mainstream HCI, but need to consider the particular characteristics of wearables. In the elicitation of user requirements, particular attention needs to be paid to developing appropriate metaphors for the devices, and not relying on existing and possibly inadequate metaphors. In the evaluation of wearables, the different possible types of use of the wearable (serial or concurrent), need to be considered when developing evaluation methodologies and measures.

Acknowledgements

The MoBIC Project (TP 1148) was supported by the Technological Initiative for Disabled and Elderly (TIDE) Programme of the European Union (DG XIII). The project partners are: BT, U.K.; F.H. Papenmeier GmbH, Schwerte, Germany; Free University of Berlin, Germany; Universität Magdeburg, Germany; Royal National Institute for the Blind U.K; University of Birmingham U.K.; University of Hertfordshire, U.K.; Uppsala University, Sweden. We would like to thank the MoBIC team and the many people who participated in the MoBIC user studies.

References

[1] Carroll, J. M., Kellogg, W.A. and Rosson, M.B. (1991). The task-artifact cycle. Carroll, J. M. (Ed.), *Designing interaction*. Cambridge: Cambridge University Press.

[2] Clark-Carter, D. D., Heyes, A. D., & Howarth, C. I. (1986). The efficiency and walking speed of visually impaired people. *Ergonomics*, *29*(6), 779 - 789.

[3] Epstein, W., Hughes, B., Schneider, S.L. and Bach-y-Rita, P. (1989). Perceptual learning in an unfamiliar modality. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 28 - 44.

[4] Heyes, A. D., Armstrong, J. D., & Willans, P. R. (1976). A comparison of heart rates during blind mobility and car driving. *Ergonomics*, *19*, 489 - 497.

[5] National Aeronautics and Space Administration (NASA)., Ames Research Center, Human Performance Research Group. *NASA Task Load Index. Version 1.0: Paper and pencil package*. Moffet Field, CA: Author.

[6] Nielsen, J., & Levy, J. (1994). Measuring usability: preference vs. performance. *Communications of the ACM*, *37*(4), 66 - 75.

[7] Peake, P. and Leonard, J.A. (1971). The use of heart rate as an index of stress in blind pedestrians. *Ergonomics*, *14*(2), 189 - 204.

[8] Petrie, H. L. (1995). User requirements for a GPS-based travel aid for blind people. In J.M. Gill and H. Petrie (Eds.), *Proceedings of the Conference on Orientation and Navigation Systems for Blind Persons*. Hatfield: University of Hertfordshire.

[9] Petrie, H., Johnson, V., Strothotte, T. Raab, A., Fritz, S., and Michel, R. (1996). MoBIC: designing a travel aid for blind and elderly people. *Journal of Navigation*, *49*(1), 45 - 52.

[10] Shingledecker, C. A. (1978). The effects of anticipation on performance and processing load in blind mobility. *Ergonomics*, *21*, 355 - 371.

[11] Spielberger, C.D. (1983). *State-Trait Anxiety Inventory for adults: Sampler set, manual, test and scoring key*. Palo Alto, CA: Mind Garden.

[12] Strothotte, T., Fritz, S., Michel, R., Raab, A., Petrie, H., Johnson, V., Reichert, L. and Shalt, A. (1996). Development of dialogue systems for a mobility aid for blind people: initial design and usability testing. *Proceedings of ASSETS �96: The Second Annual Conference on Assistive Technology*. New York: ACM Press.

[13] Available from Visionics Corporation of Minneaplois, MN, U.S.A. Homepage availabe at: http://www.visionics.com/