# The SenseCam as a Tool for Task Observation

Daragh Byrne[1], Aiden R. Doherty[1], Gareth J.F. Jones[1], Alan F. Smeaton[1], Sanna Kumpulainen[2] and Kalervo Järvelin[2],

[1]Centre for Digital Video Processing & Adaptive Information Cluster, Dublin City University, Ireland
[2]Department of Information Studies, University of Tampere, Finland

daragh.byrne@computing.dcu.ie

## ABSTRACT

The SenseCam is a passive capture wearable camera, worn around the neck and developed by Microsoft Research in the UK. When worn continuously it takes an average of 2,000 images per day. It was originally envisaged for use within the domain of Human Digital Memory to create a personal lifelog or visual recording of the wearer's life, which can be helpful as an aid to human memory. However, within this paper, we explore its applicability as a tool for use within observational and ethnographic studies. We employed the SenseCam as a tool for the collection of observational data in an empirical study, which sought to determine the information access practices of molecular medicine researchers. The affordances of the SenseCam making it appropriate for use within this domain, as well as its limitations, are discussed in the context of this study. We found that while the SenseCam, in its current form, will not offer a complete replacement of traditional observational methods, it offers a complimentary and supplementary route to the collection of observational data.

## Categories and Subject Descriptors

H.1.2 [**Models And Principles**]: User/Machine Systems - *Human factors*; H.5.m [**Information interfaces and presentation (e.g., HCI)**]:Miscellaneous.

## General Terms

Human Factors.

## Keywords

Microsoft SenseCam, task observation.

## 1. INTRODUCTION & BACKGROUND

The SenseCam is a small lightweight, wearable device that passively captures a person's day-to-day activities as a series of photos and sensor readings [10]. Photos are captured from the perspective of the owner as the device is worn around the neck (as illustrated in Figure 1), and is so oriented towards the majority of activities which the user is engaged in. Owing to its fisheye lens, anything within the view of the wearer can be captured by the SenseCam. At a minimum the SenseCam will take a new image approximately every 50 seconds, but sudden changes in the environment of the wearer as detected by onboard sensors, can trigger more frequent photo capture. The device requires no manual intervention by the user as its sensors detect and record changes in light levels, motion and ambient temperature and then determine when is appropriate to take a

photo. For example, when the wearer moves from indoors to outdoors a distinct change in light levels will be registered and photo capture will be triggered. The SenseCam takes images quite frequently (approximately once every 20 seconds) and as such collects an average of 2,000 images in a typical day. As a result a wearer can very quickly build a large and rich, extensive visual diary of their day. The battery is sufficient to allow the camera to run all day, and can be fully recharged overnight, allowing uninterrupted recording during the day.

The Microsoft SenseCam was developed in order to more fully realise Bush's 1945 MEMEX vision [5] of a world where the fallibilities of human memory are augmented by a digital surrogate. While the SenseCam remains a key tool within human digital memory research [2], its potential applications are not limited to this. Already, the SenseCam has been deployed in variety of other contexts [17] such as tourism [21][4], patient care [3][12], education [1][9] and accessibility within business [20]. Wood, Fleck and Williams have previously discussed the potential range of applications of the SenseCam: in particular they highlight its promise in the area of tourism [21]. Further to this Blighe *et al.* have explored enhancing the experience of museum visitors with take-home SenseCam images and related content [4]. Of most note are the preliminary findings of a study between Microsoft Research and Addenbrooke's hospital in Cambridge, U.K which indicates that a SenseCam may offer significant benefits to individuals with neurodegenerative memory problems [3][12]. Passive capture devices, similar to the SenseCam, have also been used in the observation of dietary habits [18]. This use suggests that the SenseCam, and by implication personal passive capture devices, have utility in more general observation tasks. However, as yet, the SenseCam has not been explored as a tool for HCI researchers nor has its suitability to the collection of observational data within situated field studies been reported.

Ethnography, task-observation and field studies are often carried out in order to better understand the social perspective of task operation in order to design systems, which more appropriately support them. Through observation, systems can incorporate an understanding of the user's work setting, social context and pattern of activities [6]. In order to gain this understanding, a great deal of time is spent observing users, usually several weeks or months, making it a very time and resource intensive activity [16]. Typically observation is a naturalist method *"relying upon material drawn from the first hand experience of a field worker in some setting"* [13]. However, it is not uncommon for mobile devices or recording devices to be used within such activities to aid the observer, minimise the amount of time and resources required, and/or to supplement periods when the observer cannot be present [6]. Audio, video and computer recording are commonly used during observational activities [7]. Handheld PCs have also been used to aid data capture during observation [19] while mobile phone based diaries have been used to collect data from participants outside of direct observation [14]. The SenseCam however, being wearable, and both intelligently and passively

capturing a visual record from the perspective of the user may offer a novel approach to observing users in context without intrusion. It additionally may offer advantages over other recording apparatus typically used and offer a route to reducing the resource commitments required for such studies.

As such, we undertook an exploration of the appropriateness of the Microsoft SenseCam within user and task observation. We explore the use of the SenseCam within task observation for situated fieldwork and ethnographic studies. The strengths and weaknesses of the SenseCam and its applicability as a novel data collection tool for such work are discussed. These positions are supported by our experiences of its use in conducting an observational study into information access in molecular medicine research. We are not aware of any similar work relating to the SenseCam.
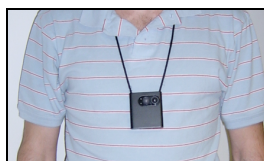


**Figure 1 SenseCam device (worn around neck)**

## 2. AFFORDANCES OF THE SENSECAM

The SenseCam offers unique affordances particularly suiting it to use within task observation activities. Unlike other recording apparatus, such as digital photo or video cameras, typically used within observation, as it is worn it proactively captures activities from the perspective of the user without intruding into their ordinary environment. Being worn, it allows a user to move around freely without impeding capture of their activities. A video camera would need to be repositioned ad-nauseum to allow such free-form activities to occur. Additionally, such recording apparatus may be intrusive into the environment of the worker, particularly in an office environment.

Ethnographic and task-observation studies normally require intensive effort, commitment and resources [13][7] and this has motivated the development of methods such as Rapid Ethnography [17] and Autoethnography [6]. The practitioner typically "shadows" or follows a participant taking extensive notes on their day-to-day activities, tasks and interactions in order to better understand the interplay between the operator and task performance. Ethnographic studies can require participants to be shadowed for weeks and possibly months on end and places an enormous burden on the practitioner to be alert and present at all times. It is conceded that while observation provides insights into the social and cultural aspects of computer use, the observer cannot be ever-present [6]. However, there is a risk that critical events can occur for the participants during these unobserved periods. While ideally, the observer would be present, this is far from feasible and so alternative methods to capture these events must be sought.

As the SenseCam is a fully automated capture system it removes the implicit need for an observer to be present at all times in certain situations. For example, as in our study, the participants may initially be shadowed for a short period of a few days following which they can be instrumented with the SenseCam for a much longer period. The SenseCam allows observation of tasks for an extended period without the time and resource commitment normally anticipated with ethnographic investigation. It additionally offers a means to capture and preserve a recording of critical events, which would otherwise not be captured without the observer being present.

An issue encountered with the use of video recordings in observational work, is the amount of information it creates and the consequent analysis time to adequately review, annotate and probe the information it collects [7]. While the SenseCam offers a recording similar to video resulting in a large volume of images, they can be much more efficiently reviewed. For example, a full day's worth of images can be played back in approximately 3 minutes [12]. Additionally, the image collection can be managed to allow more efficient review. In earlier work we have shown how a SenseCam collection can be aggregated into discrete units through a process known as event segmentation [8]. This results in a very small number of items (approximately 20 per day) which can be rapidly inspected.

Within observational studies there may be issues relating to intrusion and privacy [6]. As the SenseCam is a wholly digital observer, a review of the dataset can be performed prior to hand-over. The participants, or their employer should they have additional concerns, can therefore remove sensitive elements from the data set before providing it to the practitioner. This allows a buffer to be placed between observation and the practitioner and allows any concerns over privacy to be resolved gracefully.

Finally, there are venues where ethnographic investigations are often not feasible. These include safety critical environments or environments in which sensitive activities are being carried out or where shadowing may not be appropriate (such as in funeral homes, areas within hospitals, care homes, military, police, etc.). The SenseCam can easily be placed into such environments without the need for a "shadow" thus potentially removing concerns that may otherwise restrict ethnography in such venues. As such the SenseCam may open up new avenues in which ethnographic investigation previously would not occur.

## 3. EVALUATION

To better understand the SenseCam's applicability to observational activities we deployed it for a period within an ongoing empirical study. By comparing it with the more traditional 'shadowing' techniques, qualitative findings on its suitability within this context of use were uncovered and are reported in the following sections.

### 3.1 Evaluation Details

As part of the evaluation of the SenseCam in task observation activities, it was included in an observational study being conducted by the University of Tampere. The empirical study, which was conducted over a six-month period (2007-2008), sought to determine information access practices of molecular medicine researchers. The information access practices of these researchers are not limited to electronic means as they may uncover relevant material from printed media or through discussions with coworkers, colleagues or supervisors. The SenseCam was employed to observe the full range of information access practices in combination with software which logged the details of electronic access. The SenseCam was used to: overview and observe the practices and the patterns of knowledge acquisitions; and to preserve a record of this acquisition which could be later probed during follow up discussions in order to uncover the specifics of or the motivations for information access. There were six researchers who participated in the study. Each of the individuals was shadowed using traditional practices over periods ranging from three to eight weeks, totaling on average 24 hours observation per person. In addition to the direct observation, each of the participants wore a SenseCam for a short period. Five of the participants wore it for 10 working days, with one wearing it for two days. The period of SenseCam recording partially overlapped with those that were shadowed, in order to provide comparison between the two approaches.

## 3.2 Supporting Software

Given that an individual will collect an average of 2,000 images per day or approximately 30,000 images during the course of the evaluation study, we recognised the need for the collection of data to be made more manageable in order to enable practical use of it. As a result, prior to review we aggregated sequences of SenseCam images together into distinct events or activities e.g. walking into work, sitting at PC, talking to colleague, at lunch, etc. This process of event segmentation [8] breaks an average day's images into approximately 20 events.

To help review this collection of data, the segmented image collections were presented within an event-based browser (illustrated in Figure 2), which allowed the selection of an individual event of interest and the display of its constituent images. As these images are temporally consistent they will often 'storyboard' the progression of a task or operation. This 'storyboard' offers the practitioner a means by which specific elements of a task and its progression can be easily probed and discussed with the participant. The tool also offered a means by which discussion could be focused around particular instances or interactions and a means by which both the participant and the practitioner could better illustrate their discussion with concrete examples. As such, the SenseCam images, as a visual account of tasks a participant carried out, and their segmentation into 'events' proved very useful within post-observation discussions and/or follow-up interviews.



**Figure 2 SenseCam browser**

## 3.3 Observed Limitations of the SenseCam

The outcomes of the study, feedback from participants and practitioners all confirmed that the SenseCam offers many of the affordances as outlined previously. However, it also highlighted that its use in observational work is not without its limitations. One major drawback to its use is that there is an onus placed on the participants to wear it and capture recordings regularly and reliably during their working day. We encountered a number of difficulties in this regard as the SenseCam was prone to being forgotten by the participants! On several occasions participants forgot to switch on the device as they began work in the morning, not noticing their error perhaps for several hours. Participants occasionally removed or switched off the device in situations where privacy was required, e.g. using the bathroom. Again the SenseCam was easily forgotten, either being left behind on the desk as participants carried out their tasks or not switched back on. Others even forgot to remove the lens cap on occasion, resulting in useless blank images.

While it is less intrusive than shadowing, users are still nevertheless aware of the presence of the SenseCam and hence do not act in a completely natural manner. In a study carried out by McAtamney & Parker it was found that users are very conscious of wearing lifelogging devices [15]. We noticed a similar effect among our participants initially, although after wearing the device over a number of days the participants became much less conscious of it. Not only is the wearer of the device quite conscious of it, but also other people in the vicinity are aware of the device and feel as if they are "under surveillance". As a result of this it was necessary to spend time reassuring the workgroup of the benefits offered by the passively captured images from the SenseCam.

By implication we can assume that this feeling of 'being under surveillance' may have an impact on general social interactions with the wearer. For example, we have observed people more carefully choosing their seating to avoid being captured by the device. As fieldworkers, and by proxy the devices they employ within a study, require acceptance by participants and their colleagues [13], this may appear to be problematic. However, this feeling typically subsides after a few days once colleagues grow more accustomed to the device and interactions tend to return to normal at this point. As such, much reminding, encouragement and reinforcement is required when using the device in such situations.

As previous studies have indicated, images captured by the SenseCam are often of low quality and with the main subject prone to being obscured [11]. Our experiences in observational fieldwork with the SenseCam mirrored these findings. A large number of the images were of low quality with significant blurring or suffered from light saturation as the wearer moved from indoors to outdoors or perhaps faced a window. In poorly lit conditions they were often extremely dark or unusable as the SenseCam has no flash or infra-red photo capture capability. Additionally, the images were sometimes obscured fully or partially by the wearer's clothing. This was particularly observed in outdoor situations, where jackets and coats often covered the lens of the device.

The low image capture rate of the SenseCam caused additional problems when reviewing captured data. The SenseCam can often miss large amounts of visual change. While the onboard sensors help to mitigate against this, intelligently determining if a photo should be taken based on movement of the wearer, for example, it doesn't eliminate the issue entirely. The observation of detailed and/or expert computer-based tasks was particularly problematic. As expert users will often operate computer software using rapid iterations of keyboard shortcuts, clicks and commands, the SenseCam cannot capture the subtleties of these interactions. Its frame-rate simply does not support it. It does however offer a 'gisting' or high-level overview of the progression of such tasks but this is often insufficient, particularly as we sought to observe information access which is a predominantly computer based activity. This however could be overcome if the SenseCam were to be used in combination with a keystroke logging or screen recording tool. Furthermore, while the sequence of visual images produced by the SenseCam is usually effective in identifying actions within person-to-person interactions, its lack of audio recording makes it difficult to identify contextual factors of the activity. This can sometimes make it difficult to distinguish between a social discussion or a task-based discussion.

## 3.4 Discussion

Given the above issues it can often be difficult to gain a complete understanding of a user's activities using the SenseCam alone. It does however offer a route to collecting situated observational data, which can compliment and supplement those shadowed periods. For example, we found the SenseCam recordings allowed us to further extend and validate our findings from shadowed observation. Using shadowing the features of the tasks performed by participants were determined. The SenseCam data was then used to better gauge the frequency and criticality of these tasks and sub-tasks in unshadowed day-to-day activities. This greatly helped to better

generalise our findings and identify the significance of problems encountered in task performance.

The use of the SenseCam is however not without its drawbacks. Most of these stem from the lack of direct communication between the observer and observed participant. This communication is used in shadowing to allow the observer to gain a clear understanding of the contextual factors, which have an effect on how the task or tasks are conducted. As the SenseCam removes this communication channel, it can make it difficult to fully understand the task under observation. This can be overcome to some degree with post-observation follow-ups using the SenseCam images to probe items of interest. However, we feel strongly that having a human in the loop at the time of task execution is invaluable.

The SenseCam in its current form can only offer a supplementary source of data within the observation process and not a full replacement of shadowing techniques. As it is a prototype device, this is not to say however that with future versions the limitations mentioned above may be overcome, making it more amenable to use in this domain. One might expect that as the device develops further, the capture rate of the device will increase, and perhaps even approach that of video. We can also expect many of the image quality issues to be resolved and the addition of other sensors or peripherals for recording, such as audio or biometric information. Additions like this would make the SenseCam a much more suitable substitute for shadowing. However, further consideration with regard to reminding participants to record and to check the SenseCam is operating would still be required when deploying to participants of such studies.

## 4. CONCLUSIONS

The SenseCam offers a novel route to the collection of observational data. When worn by the user, tasks and activities are captured as a series of images from the user's perspective. The SenseCam however is not without its limitations. Most importantly, we feel that the SenseCam should not fully replace direct observation due to the understanding and knowledge transfer between observer and participant during observation. The soft communication, which is enabled during these interactions, allow for clarification of task flow and progression and insights into the user and their work context that the SenseCam simply cannot offer. While the SenseCam, in its current form, will not offer a complete replacement of shadowing and traditional observational methods, it offers a number of unique affordances and benefits within such activities.

## 5. ACKNOWLEDGMENTS

## 6. REFERENCES

[1] Barreau, D., Crystal, A., Greenberg, J., Sharma, A., Conway, M., Oberlin, J., Shoffner, M., and Seiberling, S. Augmenting Memory for Student Learning: Designing a Context-Aware Capture System for Biology Education, *Proc. of the American Society for Information Science and Technology*, 43(1) pp. 251–251, Oct 2007.

[2] Bell, G. and Gemmell, J. A Digital Life. *Scientific American*, 2007.

[3] Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B., and Wood, K. The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis. *Special issue of Neuropsychological Rehab.* 2007, 17 (4/5), pp. 582-681.

[4] Blighe, M., Sav, S., Lee, H., and O'Connor N. Mo Músaem Fíorúil: A Web-based Search and Information Service for Museum Visitors. *In Proc. ICIAR 2008*.

[5] Bush, V. As We May Think, *In The Atlantic Monthly*, 176, 1, pp. 101-108, (July 1945).

[6] Cunningham, S.J. and Jones, M. 2005. Autoethnography: a tool for practice and education. *In Proceedings of CHINZ '05*, vol. 94. ACM, New York, NY, 1-8.

[7] Diaper, D. (1989) *Task Observation for Human-Computer Interaction*. In D. Diaper (Ed.), Task Analysis for Human Computer Interaction. Ellis Horwood, Chichester, England. pp. 210-237.

[8] Doherty, A.R. and Smeaton, A.F. Automatically Segmenting Lifelog Data into Events, *Proc. WIAMIS '08*.

[9] Fleck, R. and Fitzpatrick, G. Supporting Collaborative Reflection with Passive Image Capture, *In Supplementary Proceedings of COOP'06*, pp. 41-48.

[10] Gemmell, J., Williams, L., Wood, K., Lueder, R., and Bell, G. Passive Capture and Ensuing Issues for a Personal Lifetime Store, *Proceedings of the 1st ACM workshop on continuous archival and retrieval of personal experiences* New York, USA October 2004.

[11] Gurrin, C., Smeaton, A.F., Byrne, D., O'Hare, N., Jones, G., and O'Connor, N. An Examination of a Large Visual Lifelog. *In Proc. AIRS 2008*, Harbin, China, January 2008.

[12] Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., and Wood, K. SenseCam: a Retrospective Memory Aid. *Proc. Ubicomp 2006*, pp. 177 –193.

[13] Hughes, J., King, V., Rodden, T., and Andersen, H. 1995. The role of ethnography in interactive systems design. *Interactions* 2(2), Apr. 1995, pp. 56-65.

[14] Hulkko, S., Mattelmäki, T., Virtanen, K., and Keinonen, T. 2004. Mobile probes. *In Proc. NordiCHI '04*, pp. 43-51.

[15] McAtamney, G. and Parker, G. An Examination of the Effects of a Wearable Display on Informal Face-to-Face Communication*, In Proc. SIGCHI 2006*.

[16] Millen, D. R. Rapid ethnography: time deepening strategies for HCI field research. *In Proc. of DIS'00*, Aug. 2000. pp. 280-286.

[17] Other Applications for SenseCam. Retrieved from: http://research.microsoft.com/sendev/projects/sensecam/applications.htm .

[18] Reddy, S., Parker, A., Hyman, J., Burke, J., Estrin, D., and Hansen, M. 2007. Image browsing, processing, and clustering for participatory sensing: lessons from a DietSense prototype. *In Proc. EmNets '07*. pp. 13-17.

[19] Spinuzzi, C. 2003. Using a handheld PC to collect and analyze observational data. *In Proc. SIGDOC* '03. 73-79.

[20] Tjoa A., Andjomshoaa A., and Karim S. Exploiting SenseCam for Helping the Blind in Business Negotiations, Computers Helping People with Special Needs, *Springer*, pp. 1147 – 1154, 2006.

[21] Wood, K., Fleck, R. and Williams, L. Playing with SenseCam. *In Proc. Playing with Sensors (W3) at UbiComp 2004*, Nottingham, UK.