M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 468
To Appear: Proc. of the Intl. Symposium on Wearable Computers, 1998

StartleCam: A Cybernetic Wearable Camera

Jennifer Healey Rosalind W. Picard
MIT Media Laboratory
20 Ames St.
Cambridge, MA 01239
{fenn, picard}@media.mit.edu

Abstract

StartleCam is a wearable video camera, computer, and sensing system, which enables the camera to be controlled via both conscious and preconscious events involving the wearer. Traditionally, a wearer consciously hits record on the video camera, or runs a computer script to trigger the camera according to some pre-specified frequency. The system described here offers an additional option: images are saved by the system when it detects certain events of supposed interest to the wearer. The implementation described here aims to capture events that are likely to get the user's attention and to be remembered. Attention and memory are highly correlated with what psychologists call arousal level, and the latter is often signaled by skin conductivity changes; consequently, StartleCam monitors the wearer's skin conductivity. StartleCam looks for patterns indicative of a "startle response" in the skin conductivity signal. When this response is detected, a buffer of digital images, recently captured by the wearer's digital camera, is downloaded and optionally transmitted wirelessly to a webserver. This selective storage of digital images creates a "flashbulb" memory archive for the wearable which aims to mimic the wearer's own selective memory response. Using a startle detection filter, the StartleCam system has been demonstrated to work on several wearers in both indoor and outdoor ambulatory environments.

1 Introduction: Cybernetic Control of Wearable Computers

We use the expression "cybernetic control" to describe the control loops formed by the nervous system, the brain, and the wearable mechanical-electrical communication system. The StartleCam control links these two feedback systems, using an autonomic nervous system response–the skin conductivity startle response–to automatically control the electro-mechanical wearable computer and communicate selective information to the wearer's local digital archives or a remote server on the Internet.

This control loop might also be considered "cyborgian" as the term cyborg denotes man in quasi-symbiotic union with electro-mechanical homeostatic control systems. The concept of a cyborg was originally envisioned for astronauts to be able to regulate their homeostatic environment without conscious attention in the hostile environment of space[Cly95]. However, the idea is equally relevant to people on earth attempting to regulate their homeostasis against the stress created by information overload. By using physiological cues, the wearer's computer can react and respond in real time to unexpected events. In this aspect, StartleCam has an advantage over software agents that use a priori belief models of the user preferences [Mae94], although agents may someday use these cues to help train their belief models.

The startle response has been linked to reactions of sudden fright [LeD94], anticipation of bad results [Dam94], and stressful situations [Hel78] [Lev92] that are deviations from homeostasis [LG88]. It is also often referred to as the orienting response, because any interruption in attention can generate it, not jut the extreme startle reaction. By saving images when the startle response is detected, the StartleCam application models the wearer's own capacity for selective memory, according to a theory that memories are formed when survival is threatened [Dam94]. This effect is what is responsible for the the so-called "flashbulb memory [BK77]" where extremely arousing events seem to be highlighted and stored with unusual clarity. StartleCam offers three different modes of operation, direct control, automatic logging, and image series capture. The camera can be controlled directly by the wearer, through the input device, to record any data of specific interest as snapshot or an image series. When the startle response is detected, a composite image of recent events can be created by saving several frames from the video camera in rapid succession as shown in Figure 8. By setting a high threshold for the startle detector, StartleCam will record only the most arousing or threatening events. This mode of operation would be most useful for an application such a "SafetyCam" in which images of threatening events are transmitted to secure websites of the wearer's SafetyNet of friends and family[Man97]. In the opposite mode, the camera can be set to automatically record images at a set frequency when very few responses have been detected from the wearer, indicating that their attention level has dropped. This mode of operation is useful when the value of a lecture, a meeting or a contact is not immediately apparent and our natural selective memory system designed to survive immediate threats fails us.

StartleCam's cybernetic control loop allows the wearer to control applications on the wearable computer with minimal effort. This allows the wearer to offload tasks for which computers are well suited. For example, human memory tends to be inaccurate for recalling the exact details of events[BK77], and memories tend to change over time, but a computer can easily store a video archive with time stamps to aid recall. The vision of the cyborg has the computer cooperating with the human to manage cognitive load and mediate the flow of and storage of information; however, direct control should never be taken from the user. StartleCam allows the user to directly add or delete images from the video archive and start or stop an automatic process of capturing video images at regular time intervals throughout the day.

2 The StartleCam System

StartleCam is a wearable computer system with a digital video camera, physiological sensors and a wireless connection to the Internet. The wearable computer used in this prototype is the MIT Media Lab's Lizzy design[1] using a 100MHz 486 processor with a two gigabyte hard drive [Sta95]. The Lizzy computer is used to run the real-time signal processing algorithm to detect the startle response, to hold a rotating buffer of video images from the digital camera, and to control the flow of video information to the hard disk or the Internet.

The digital video camera shown here is a black and white Connectix QuickCam, which can capture up to 15 frames per second of 320 by 240 pixel images in 6-bit grey. The camera has a fixed focus, from eighteen inches to infinity, and automatic white-balance calibration. A compression script stores images from the QuickCam in a JPEG format, which requires only 1-3 kilobytes of memory per image.

Images from the camera are stored in a virtual buffer until the StartleCam detection algorithm is triggered. This system maintains a five-image buffer of images recorded at one frame per second. When a trigger is detected the entire buffer is saved as a single image and can be both downloaded to the hard drive or sent over the Internet to a remote server.

To transmit information back to the Internet, two options are available with this system: a two megabits per second wireless ethernet connection using Digital's RoamAbout/DS and a 14.4 kbps CDPD connection using a Sierra wireless modem. The RoamAbout connection is preferred, but it is unfortunately limited to indoor applications. The slower CDPD connection is necessary for outdoor monitoring.

The skin conductance is measured by applying a small voltage to two electrodes and measuring the resultant current conduction of the skin. This analog signal is then sampled at 20 samples per second using the ProComp system from Thought Technologies. Although the electrodes shown here are placed on the index and middle fingers of the hand (as shown in Figure 1) the eccrine sweat glands that produce this response cover most of the body, allowing sensors to be located comfortably in a number of places. The palms and soles are preferred only because they have the highest concentration of eccrine glands[SF90]. Electrodes measuring these areas could be worn as a pair of rings or embedded into the insoles in shoes [PH97].

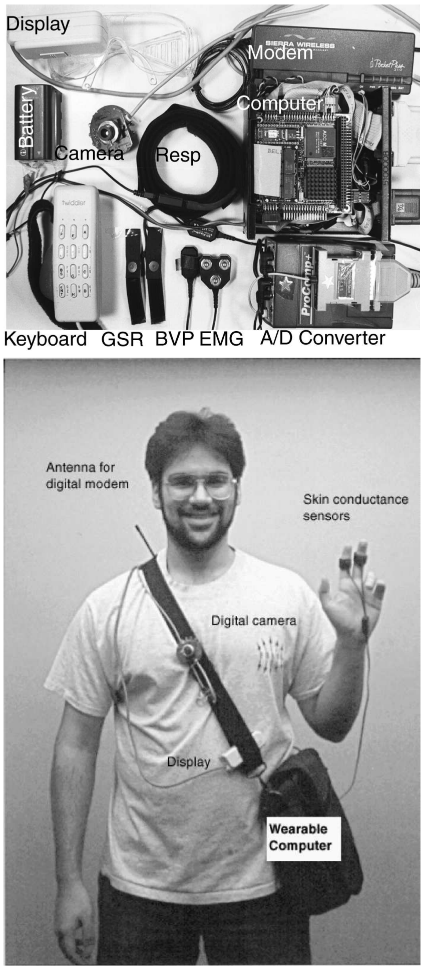

Figure 1: The StartleCam system consists of a skin conductivity sensor (GSR) which is sampled by an analog to digital converter attached to a wearable computer. A digital camera and digital modem are also attached to the computer. Images are captured by the digital camera and stored in a buffer in memory. When the computer algorithm detects a startle response, the buffer of images is downloaded and transmitted wirelessly back to the Internet. Figure (a) shows the details of the system and Figure (b) shows the system as worn with skin conductivity sensors on the hand (Alternatively, the skin conductance sensors can be placed on the foot.)

3 The Skin Conductivity Response

The skin conductivity startle response is one of the most robust and well studied physiological responses. It is caused by sympathetic nervous system activation, which changes the levels of sweat in the eccrine sweat glands and has been shown to be linked to measures of emotion, arousal, and attention. The eccrine sweat glands on the palms of the

[1] http://wearables.www.media.mit.edu/projects/wearables/ has up to date information as well as a picture of this system.

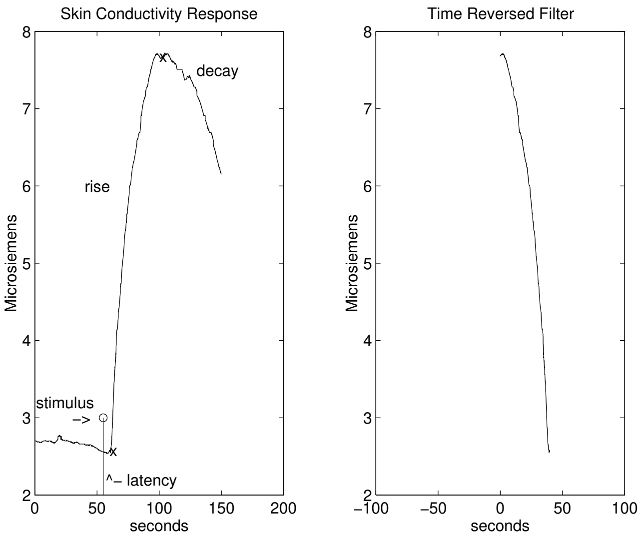

Figure 2: The matched filter is derived from a typical skin conductivity response. Figure (a) shows a response to an audio stimulus. The three phases of the response, a latency following the onset of the stimulus, a sharp rise and a consequent decay are labeled. Figure (b) shows the filter, a time reversed version of the rising edge of the response.

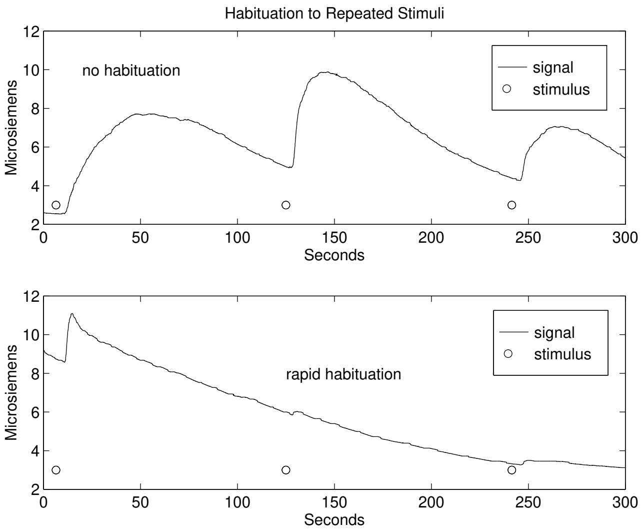

Figure 3: Habituation to a repeated stimulus occurs at different rates for different subjects. Figure (a) shows a subject with no habituation during the experiment, all the three stimuli elicit a response. Figure (b) shows, a subject with rapid habituation, only the first stimulus elicits a response.

hands and the soles of the feet are particularly responsive to emotional activation, and only minimally responsible for thermoregulation[SF90].

Ionic sweat is more conductive than dry skin, which causes an increase in bulk conductivity, proportional to the amount the glands have filled[Llo59]. An example of a measurement of the skin conductivity response is shown in Figure 2 where seimens is the unit of measurement for conductivity, equivalent to $\frac{1}{Ohms}$. The typical startle response has three components, as shown, a period of latency following the stimulus (approximately one to three seconds in duration), a rising arm, and a decay.

The magnitude of the response is related to the magnitude of the stimulus. Stronger stimuli give stronger responses[SF90]. Responses can be triggered by events including turning on a light, hearing a sound or anticipating a loss in a card game [Dam94]. However, these responses will be smaller than those caused when the wearer's survival is threatened. By setting two thresholds in the detection algorithm, StartleCam can respond differently to these two classes of events, perhaps only storing events that trigger small responses in our video archives and transmitting to our SafetyNet only those events that trigger large responses.

Although well correlated with emotional events in studies, the skin conductance response is still not entirely predictable. Even for stimuli which are supposed to be universally startling, such as white noise bursts, some subjects will show a strong response and others will habituate rapidly, not responding when the stimulus is repeated[SF90]. Examples of these reactions to the same stimulus are shown in Figure 3. The same subject may show either reaction on variables which are not fully understood, but which may depend on the time of day and their stress level. It may also happen that a wearer exhibits a startle response when there is no apparent stimulus[SF90]. Although the frequency of unstimulated responses may be an indicator of overall stress level in the wearer, they are not indicators of particularly noteworthy events. This would cause the StartleCam system to fail with a false negative result even when the detection algorithm works perfectly. In a real world application, StartleCam would detect naturally elicited responses.

4 The Detection Algorithm

The detection algorithm processes the incoming signal from the skin conductance sensor using convolution with a matched filter, the first forward difference and a threshold comparison. The filter was chosen to match the rising arm of the startle response, as shown in Figure 2. In psychophysiological studies, the rising arm of the response has proved to be both more significant and less dependent on environmental variables than the decay [SF90]. This rising arm filter is shorter than one capturing the entire response and reduces the calculations required for convolution sum in the detection algorithm. Using this filter also improves detection when the decay of the response is corrupted by a second startle. The digital filter, $f[n]$, of length l (for the filter shown $l = 800$), used for detection is a time-reversed rising edge of a typical startle response. The matched filter can be tailored to best represent the response of individual users, however the shape of the response is similar for many people and in this paper the typical response filter in Figure 2 was used for all subjects.

To detect the startle response, the incoming signal from the skin conductivity sensor, $s[n]$, is convolved with the

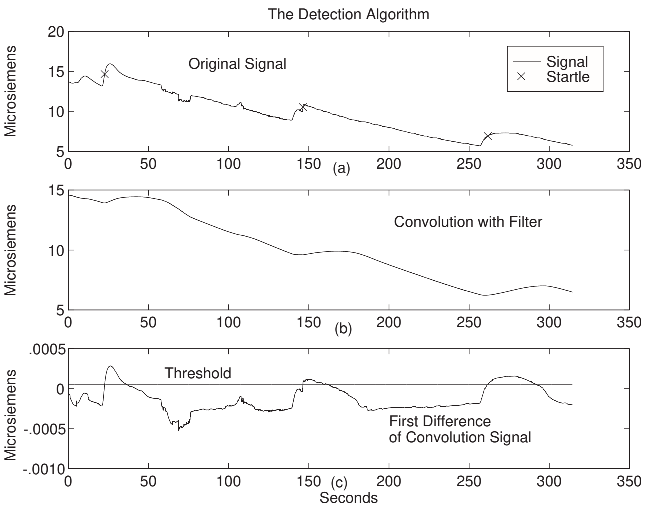

Figure 4: The Startlecam algorithm consists of three steps, convolution with a filter, taking the first forward difference and applying a threshold. Figure (a) shows the raw signal sampled from the skin conductivity sensor with x's marking where a startle was detected, Figure (b) shows the same signal after convolution with the filter, the signal is smoothed and but the long rising edges are preserved. Figure (c) shows the signal after the first forward difference is taken from the signal in Figure (b). The startle response is detected at the point where the signal in Figure (c) crosses above the threshold..

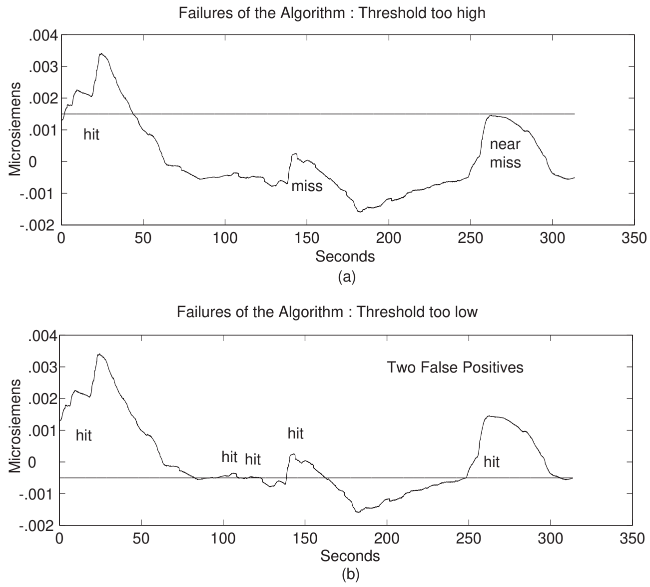

Figure 5: Varying the threshold changes the accuracy of the algorithm. Figure (a) shows missed responses from a threshold that is too high. Figure (b) shows false positive responses from a threshold that is too close to baseline.

time-reversed matched filter $f[n]$, resulting in the convolution sum, $c[n]$, according to the equation:

$$c[n] = \frac{1}{N} \sum_{k=0}^{l} s[k] f[n-k], \tag{1}$$

where N is the normalization factor for the matched filter:

$$N = \sum_{n=0}^{l} f[n]. \tag{2}$$

The first forward difference, d[n], is then calculated for every point of the convolution sum:

$$d[n] = c[n] - c[n-1]. \tag{3}$$

When the value of the difference vector, $d[n]$, changes from a value below a set threshold to a value above that threshold, the algorithm declares that a startle has been detected. A second startle will not be detected until the value of the difference vector drops below and then above the threshold. Also, the algorithm will not detect a startle in the first ten seconds of processing, allowing time for the buffer of video images to fill and the leading artifacts of the convolution sum to be cleared from the convolution buffer.

5 Experimental Validation

The startle respone varies both between individuals, as shown in Figure 3, and for each individual at different times, as shown in Figure 7. The robustness of the startle detection algorithm was tested on 27 responses representing both differences between individuals and differences for an individual at different times. In a startle eliciting experiment, subjects were seated, wearing a skin conductivity sensor with electrodes on the index and middle finger and asked to listen to three stimuli (50 millisecond audio bursts of 95 dB white noise). These stimuli were designed to elicit three startle responses. Of the eleven subjects who participated in the study, nine subjects responsed to the stimulus, but two habituated rapidly as shown in Figure 3 and were eliminated.

As the threshold of the startle detection algorithm was varied the number of missed responses and the number of falsely detected responses varied. If the threshold is set too high, some of the startle responses are not detected, and if the threshold is set too low, then both false positives and false negatives result as the threshold passes below the skin conductance baseline, as shown in Figure 4. For this pool of responses, a zero error rate was found with a threshold of 0.0005 microseimens. However, using this same method, the threshold could be customized for an individual user from multiple responses over time.

6 The Robustness of the Algorithm for Alternate Sensor Placements

To make such a system truly wearable it should be unobtrusive and fit comfortably into normally worn clothing, jewelry, and accessories. In earlier work, we explored the

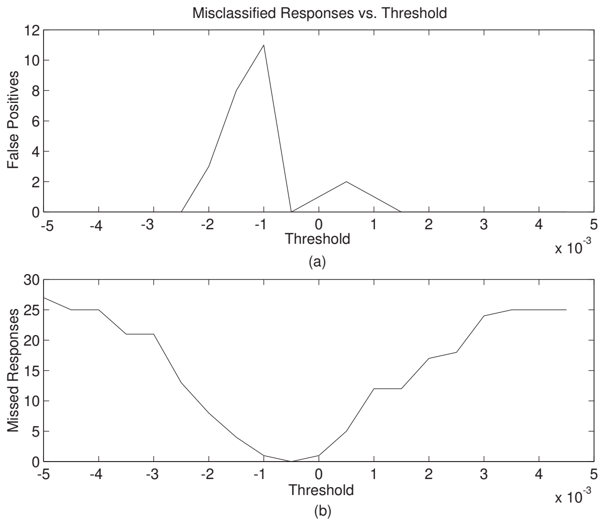

Figure 6: Adjusting the threshold creates two kinds of errors, false positives when the threshold is too low (but intersecting the signal) and missed responses when the signal is too high or low. Figure (a) is a record of the number of false positives for the startle responses as the threshold is varied, Figure (b) is a record of the responses missed by the algorithm.

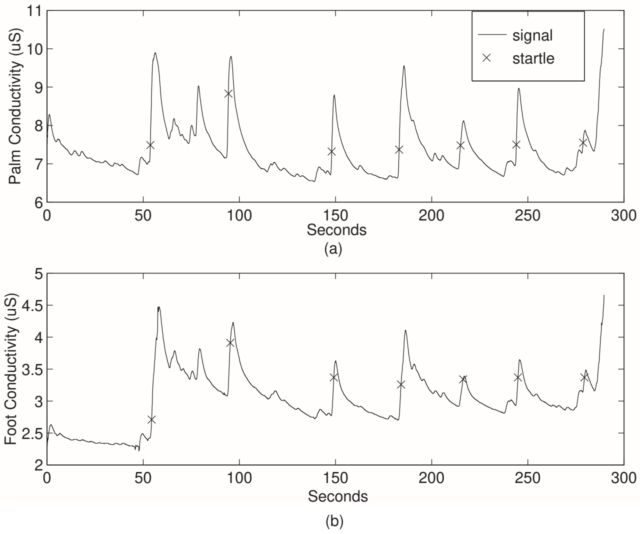

Figure 7: The StartleCam algorithm can detect responses in skin conductivity signals from placement on either the hand or the foot. Figure (a) shows the detected startles on a signal from the hand and Figure (b) shows the algorithm working on a signal taken from the foot at the same time.

option of placing the electrodes on the toes and measuring skin conductivity across the sole of the foot [PH97]. We evaluated the startle detection algorithm on these measurements, to see if the startle response could be detected as easily from the foot as from the hand. An example of the startle detection algorithm working on simultaneous data from the foot and the hand is shown in Figure 7. The algorithm detects startles in the foot's signal that match the startles detected from the hand's signal. Even more placement options should yield similar results, since the eccrine sweat glands that produce this response cover the entire body. Because the response uses a bulk measurement conductivity, measuring a large area of skin can compensate to the smaller concentration of glands in areas other than the palms of the hands and the soles of the feet. Other researchers in our group are currently exploring skin conductance readings taken from the large area of the back in electrodes embedded into a shirt worn by a conductor of a symphony[MP98].

The skin conductance response is one of the most reliable physiological signals[SF90] and should be the easiest measurement to incorporate into clothing. The startle detection algorithm detects only sharp changes in the skin's conductivity response and is not triggered by slow changes in baseline. Therefore, sweat gland activation for thermal regulation or baseline drift from deterioration in the conductance of the electrodes over time will not generate false startle responses.

7 Reducing Information Overload

Advances in technology enable us to store more information about our world, so much that many people suffer from information overload. Although we have the capacity to store a digital video record of a person's entire day using the QuickCam, such a record would take hours to review. StartleCam offers an alternative that selectively stores a series of images when the startle response is detected, indicating that these images are of possible interest to the wearer[LeD94]. In this system, the digital camera captures an image once every second and stores those images in a rotating buffer of five images. When the startle response is detected, this indicates that a startling event probably occurred in the last one to three seconds (due to the latency in the response as described in the last section). The startle trigger causes the entire buffer of images from the last five seconds to be downloaded. This can then both be transmitted via the wireless link back to a remote location or simply saved locally on the wearable's hard disk. Figure 8 show a series of images captures in the StartleCam buffer when the wearer reacted to an sudden question from a friend at a nearby workstation. The series shows the images captured as the wearer turned from her computer to look at her friend. This series of images might be considered more interesting than screen shots of the wearer's own terminal which are more or less identical. By storing only selected events at high resolution the StartleCam can help manage the wearer's information load.

8 Applications and Future Work

StartleCam reacts to the user's real-time psychophysiological reactions with a selective camera recording device. Beyond SafetyCam and selective video memory, algorithms incorporating the startle detector are being developed to measure the wearer's stress level. The control system will allow many agents to cooperate with the wearer to manage other aspects of their task load. If you choose, the computer could respond to detecting a high stress state by

Figure 8: StartleCam creates a composite image from a buffer of digital images recently captured by the wearer's camera. In this sequence, the wearer was surprised by a question from a friend. The sequence shows images captured as the wearer turned to address her friend.

filtering your less urgent e-mail or playing your favorite relaxing song. The computer could also respond to a state of low arousal by reminding you of important long term goals or presenting you with exciting news stories in your areas of interest. Over time, the computer could learn your preferences based on context, for example the computer could remember that last time you were in a high stress state at 2AM and had a paper deadline approaching, you preferred to have upbeat music played rather than relaxing music but still wanted all e-mail put on hold, except messages from your co-author. By recognizing user stress level and applying user preferences, the computer can help the wearer manage information at a level that is optimal for performance where the wearer is neither bored nor over stimulated[Kah73].

If the wearer were willing to share his responses, the data from StartleCam could be used to aid the designers of computers, cars, roads and household products. The skin conductance response is correlated with arousal, attention and effort. By reviewing data from different wearers performing tasks with StartleCam, designers could see which aspects of tasks were most difficult for wearers. This information could help researchers find better ways to improve daily chores, from attempting to diaper a baby, to creating safer roadways which were not as taxing on drivers' vigilance[Hel78].

The StartleCam system could also be triggered by physiological measurements other than the skin conductivity response. With the same system shown in Figure 1, the wearer's heart rate, respiration rate and muscle activity can be monitored as well. Algorithms could be developed to trigger the camera to activate when the wearer performed an identifiable physical task, such as standing up or when the wearer interacted with other objects which triggered responses in the wearable. Although StartleCam is only triggered by patterns in one physiological signal now, it could potentially learn more complex patterns that are a function not only of the wearer's physiology, but also of context and activity. This kind of monitoring could also augment current remote medical monitoring systems.

9 Conclusions

In the future, as our wearable computers "get to know" us better they could be empowered to help us in the way we would want them to, taking notes for us when we were not paying attention, sending a warning to our loved ones when we were in danger, adapting our flow of information to provide more when we are ready for it and less when our human buffers are full and we show signs of stress. This paper describes a wearable system, including a startle detection algorithm, which is robust across several wearers, and which enables the wearable to automatically respond to events of potential interest to the wearer. Several methods of controlling the flow of video information recorded by StartleCam have been demonstrated including transmitting images surrounding highly arousing events back to a remote server in the Internet as a safety device and the automatic logging of uneventful images when the system notices an unusually low number of responses. Working toward the vision of the cyborg as man and machine working in automatic cooperation, the physiological control of the StartleCam allows humans to offload tasks that are better handled by computers and focus on tasks that are creative and engaging.

Acknowledgments

We gratefully acknowledge BT PLC., NEC, Ricoh, and the Media Lab's Things That Think consortium for their support of this research. We would like to acknowledge Fernando Padilla for writing the scripts to implement the virtual buffer for the digital images and Justin Seger for installing the digital cameras and the wireless internet devices on the wearable system. We would like to acknowledge Thad Starner for the design of the base wearable system and Steve Mann for the idea of a wireless wearable camera-computer system that transmits digital images to a web page.

References

[BK77] Roger Brown and James Kulik. Flashbulb memories. *Cognition*, 5:73–99, 1977.

[Cly95] Manfred Clynes. Cyborg 2: Sentic space travel. In Chris Hables Gray, editor, *The Cyborg Handbook*, pages 35–42. Routledge, New York, 1995.

[Dam94] A. R. Damasio. *Descartes' Error: Emotion, Reason, and the Human Brain.* Gosset/Putnam Press, New York, NY, 1994.

[Hel78] M. Helander. Applicability of drivers' electrodermal response to the design of the traffic environment. *Journal of Applied Psychology*, 63(4):481–488, 1978.

[Kah73] Daniel Kahneman. Arousal and attention. In *Attention and Effort*, pages 28–49. Prentice-Hall, Englewood Cliffs, N.J., 1973.

[LeD94] J. E. LeDoux. Emotion, memory and the brain. *Scientific American*, pages 50–57, June 1994.

[Lev92] R. W. Levenson. Autonomic nervous system differences among emotions. *American Psychological Society*, 3(1):23–27, Jan. 1992.

[LG88] George P. Chrousos D. Lynn Loriaux and Philip W. Gold. *Mechanisms of Physical and Emotional Stress*, volume 245 of *Advances in Medicine and Biology*, section 1, pages 4–7. Plenum Press, New York, first edition, 11January 1988.

[Llo59] D. P. C. Lloyd. Response of chronergically innervated sweat glands to adrenaline and noradrenaline. *Nature*, (4682), July 25 1959.

[Mae94] P. Maes. Agents that reduce work and information overload. *Communications of the ACM*, 37(7):31–40, July 1994.

[Man97] Steve Mann. Wearable computing: A first step toward personal imaging. *Computer*, pages 25–31, February 1997.

[MP98] Teresa Marrin and Rosalind Picard. A methodology for mapping gestures to music using physiological signals. *Submitted to: International Computer Music Conference*, 1998.

[PH97] R. W. Picard and J. Healey. Affective wearables. *Personal Technologies*, (1):231–240, 1997.

[SF90] Michael E. Dawson Anne M. Schell and Diane L. Filion. The electrodermal system. In Cacioppo and Tassinary, editors, *Principles of Psychophysiology: Physical, social and inferential elements*, chapter 10, pages 295–324. Cambridge University Press, Cambridge, first edition, 1990.

[Sta95] T. E. Starner. Wearable computing. Perceptual Computing Group, Media Lab 318, MIT, Cambridge, MA, 1995.