

Subtle Gaze-Dependent Techniques for Visualising Display Changes in Multi-Display Environments

Jakub Dostal

School of Computer Science
University of St Andrews, UK
jd67@st-andrews.ac.uk

Per Ola Kristensson

School of Computer Science
University of St Andrews, UK
pok@st-andrews.ac.uk

Aaron Quigley

School of Computer Science
University of St Andrews, UK
aquigley@st-andrews.ac.uk

ABSTRACT

This paper explores techniques for visualising display changes in multi-display environments. We present four subtle gaze-dependent techniques for visualising change on unattended displays called FreezeFrame, PixMap, WindowMap and Aura. To enable the techniques to be directly deployed to workstations, we also present a system that automatically identifies the user's eyes using computer vision and a set of web cameras mounted on the displays. An evaluation confirms this system can detect which display the user is attending to with high accuracy. We studied the efficacy of the visualisation techniques in a five-day case study with a working professional. This individual used our system eight hours per day for five consecutive days. The results of the study show that the participant found the system and the techniques useful, subtle, calm and non-intrusive. We conclude by discussing the challenges in evaluating intelligent subtle interaction techniques using traditional experimental paradigms.

Author Keywords

multi-display environments; computer vision; subtle interaction; attentive user interfaces; notifications

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces—Windowing systems

INTRODUCTION

Modern computer workstation setups regularly include multiple displays in various configurations. With such multi-monitor or multi-display setups we have reached a stage where we have more display real-estate available than we are able to comfortably attend to. While the benefits of large or multi-display setups have been demonstrated in several studies (e.g. [2, 4, 7, 18]), it has also been suggested that this increase in display space will lead to usability problems [7], window management difficulties [4] and issues related to information overload [13].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IUI'13, March 19–22, 2013, Santa Monica, CA, USA.

Copyright 2013 ACM 978-1-4503-1965-2/13/03...\$15.00.

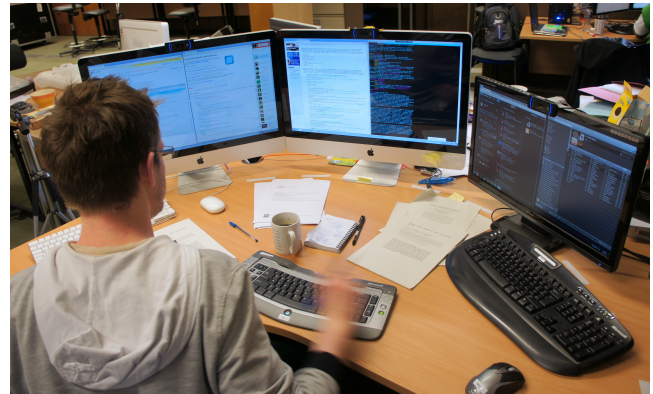


Figure 1. A photograph of the display setup used in our study. The two 27-inch displays are bezel to bezel, which translates to approximately 45° angular difference at 76.5 cm distance from the user.

Another potential issue is change blindness: users' inability to detect significant visual display changes when there is a disruption in continuity such as a brief flicker or a shift in visual focus. However, the effects of change blindness in multi-display environments have not been extensively studied in the literature. In one study, DiVita and colleagues [8] report that change blindness was a significant factor for operators managing critical events using multi-display command and control systems with unattended displays.

In general, the increased display real-estate afforded by multi-display setups means that users are unable to attend to all of it at once. In particular, this point is reached when the total display area is so large that it does not fit within the user's field of vision. In this case, the user has to substantially turn their head to see different parts of the display environment. This situation eventually arises when the number of displays, or the distance between the displays, increases. For example, it is likely to occur when users are working with three displays aligned bezel to bezel (e.g. Figure 1). When the user is only able to observe part of the multi-display environment, changes occurring on the unattended displays are difficult to track.

Attentive user interfaces [20, 21] and distraction reduction techniques [17] have been extensively studied previously using a variety of context-sensitive eye- and gaze-tracking technologies. In contrast, *inattention*, and specifically technologies that track visual change on unattended displays, has received considerably less attention [21].

One example in the literature is Ashdown and Sato's [1] multi-display system, which uses head tracking to reposition the mouse pointer to the display the user attends to. Another example is the work by Kern et al. [14] on Gazemarks: a visual placeholder that indicates the user's last fixation on a display. Kern et al. [14] reported that the Gazemarks technique resulted in faster completion times in a map navigation task in a vehicle driving scenario. Finally, Bi and Balakrishnan [4] suggest further investigation of awareness of peripheral applications with large, single and dual desktop setups as a fruitful avenue for future research. However, research on change blindness in multi-display environments has shown that making users aware of changes in unattended displays is a significant challenge [8].

In the remainder of this paper we first contribute four visualisation techniques for assisting users in perceiving and tracking display changes in multi-display environments. Second, we present an accurate markerless system that is able to track the attended display using off-the-shelf web cameras and an evaluation of its performance and accuracy. Third, we present the results of a longitudinal case study in which a working professional used our system for 39.25 hours as part of his regular work activities. Finally, we discuss the challenges in evaluating subtle intelligent visualisation techniques.

APPROACH

We have designed four subtle gaze-dependent techniques for visualising display changes that we call FreezeFrame, PixMap, WindowMap and Aura. One can view our contribution within the attentive user interface framework proposed by Vertegaal et al [21]. With respect to that framework, we are attempting to sense and communicate changes due to *inattention*.

Our techniques are designed to be used within an existing work context of the user. As a consequence, it is important that the techniques do not distract the user from their primary task. Therefore, our techniques are designed as *calm technologies*, as proposed by Weiser and Brown [23]. A calm technology "moves easily from the periphery of our attention, to the center, and back" where "the periphery is informing without overburdening" [23].

To be able to implement and deploy the non-intrusive visualisation techniques in an actual work environment, we have developed a system that is able to track the display the user is attending to and, by extension, the displays that are unattended. Importantly, our system is markerless and only relies on off-the-shelf web cameras for detecting users' gaze. This allows the system to be readily deployed in existing workstation setups. Our experimental results demonstrate that our system can accurately detect which display the user is attending.

As a subtle interface, our system does not naturally lend itself to a traditional evaluation approach, such as a short controlled experiment. Therefore, we instead carried out a qualitative longitudinal study to better understand the finer points and nuances of using visualisations as part of an ordinary multi-display environment. We recruited a working professional

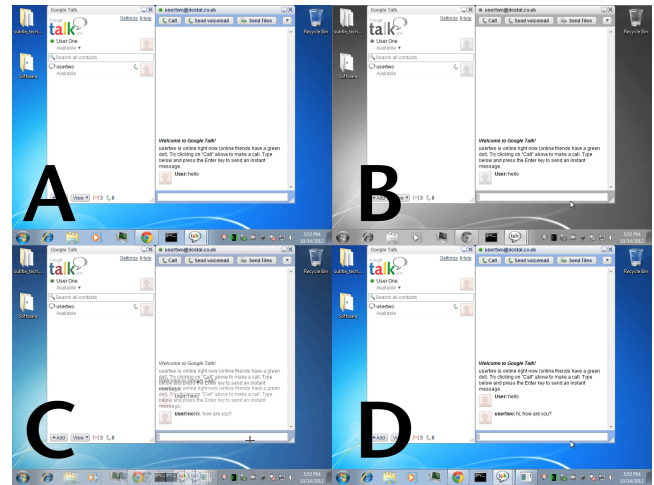


Figure 2. FreezeFrame reduces distractions by hiding visual change until the user shifts their visual focus back to the display. *A* shows the last frame before the user shifts visual focus away from this display. *B* shows a static frame of the unattended display before and during any visual change. *C* illustrates the dissolving of the old frame into the current display state. *D* shows the current display state.

who used our system during an entire workweek as part of his regular work activities. We recorded 39.25 hours of data in total. Our study shows that three out of four techniques highlighted visual changes in a subtle and non-intrusive manner. Additionally, the techniques reduced distractions and the participant found them useful when working in his regular multi-display setup.

In our study, distraction reduction was one of the main aims and benefits of using the techniques. In fact, the participant reported that before taking part in the study, he was unaware of how often he was distracted until he started using our techniques. This finding complements research into change blindness that suggests that people are not fully aware of their inability to perceive visual changes [3, 8]. Based on our study, we conjecture people may be unaware of the extent to which they get distracted by visual display changes in their environment. However, more work is required to firmly establish whether this hypothesis holds for other scenarios.

SUBTLE VISUALISATION TECHNIQUES

To assist users in noticing changes in multi-display environments, we have implemented four different subtle visualisation techniques for tracking visual changes on unattended displays. Each technique visualises display changes in a different manner. The term *visual focus* refers to an approximate direction of gaze (defined as the foveal field of vision). As gaze is a good approximation of the direction of attention [21], we use visual focus to approximate the direction of gaze. We will use the term *frame* to refer to a screenshot of the contents on a display at a specific point in time. All figures demonstrating the techniques use an instant message chat (specifically Google Talk) as an example scenario. The change that occurs while the user's visual focus is elsewhere is in all cases a new instant message appearing in the chat window.

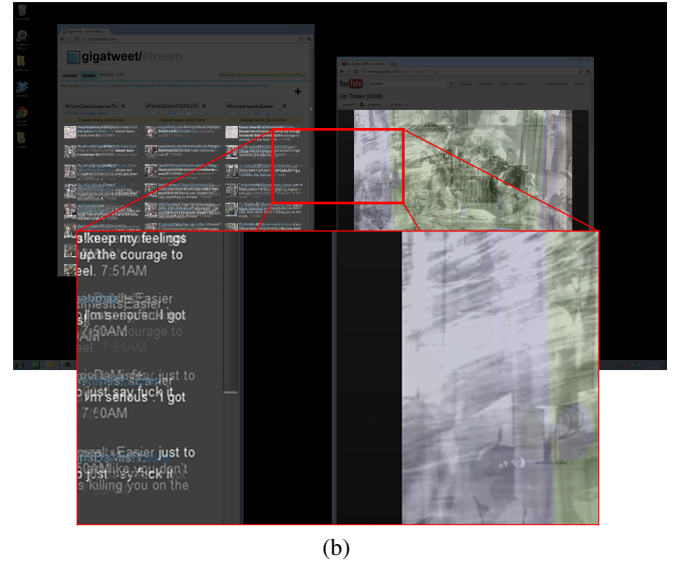
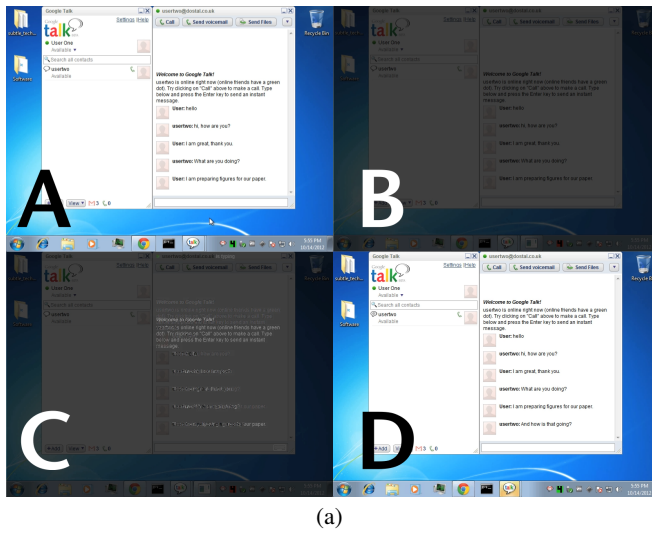


Figure 3. An illustration of PixMap. Pixmap highlights visual changes over time on a pixel level. *A* shows the last frame before the user switches visual focus away from this display. *B* shows a frame of the unattended display before any visual change takes place. *C* illustrates how the technique visualises the change (a new instant message in the chat window) when it happens by brightening the pixels that changed in the chat window. *D* shows the current display state after the technique has been dissolved. Figure 3b shows a detailed example of change visualised as bright pixels.

FreezeFrame

FreezeFrame is the simplest technique, which hides visual changes on an unattended display until the user's visual focus shifts towards it again. FreezeFrame works as follows. Consider a user working with two displays (ONE and TWO). The user starts by attending display ONE. When the user switches visual focus from display ONE to display TWO, our gaze-detection system notices this focus shift. FreezeFrame then captures a frame of display ONE at the time the user shifted their visual focus to display TWO. This frame is then displayed on display ONE as a black and white static image. No visual change is then shown on display ONE until the user switches their visual focus from display TWO back to display ONE. When the user attends to display ONE again, the frozen black and white frame on display ONE dissolves, blending smoothly into the current state of display ONE over the next 1–2 seconds. Figure 2 and Video Figure 1.1 illustrate FreezeFrame.

PixMap

The PixMap technique is a temporal heatmap visualisation that shows the user the changes that occurred while the display was unattended. Following the scenario we previously described for FreezeFrame, when the user shifts their visual focus from display ONE to display TWO, display ONE is darkened to reduce distractions. While the user's visual focus is directed elsewhere, a frame (a screenshot of the contents of display ONE) is captured at regular intervals, approximately six times per second. A difference in pixel values between the two most recently captured frames is computed for the entire display and the change is visualised on display ONE. This continues until the user's visual focus returns to display ONE, at which point the visualisation is dissolved into the current

desktop state of display ONE over the next 1–2 seconds. Figure 3 and Video Figure 1.2 illustrate PixMap.

We use the following formula to compute the value of pixels:

$$V_{new} = (V_{previous} \times Decay) + (V_{diff} \times Intensity)$$

where V_{new} is the new pixel value, $V_{previous}$ is the pixel value of the corresponding pixel in the previous iteration of the heat map. V_{diff} is the intensity of change between the last frame and the current frame measured as the difference in RGB values for the corresponding pixels. $Decay$ is a fraction denoting how quickly the current value should fade over time, and $Intensity$ defines how much of the intensity of V_{diff} will be added to the heat map.

The decay and intensity are empirically determined parameters. We set them as follows after testing the visual effect of different values:

$$Decay = 0.01, Intensity = 0.5.$$

WindowMap

WindowMap is a variation of the PixMap technique. Similar to PixMap, WindowMap also visualises changes that have occurred on an unattended display. However, instead of visualisation changes at the pixel level, WindowMap shows changes at the application-window level. WindowMap works identically to PixMap except that the amount of change in the unattended display is computed for application-window areas on the display instead of individual pixels. Figure 4 and Video Figure 1.3 illustrate WindowMap.

Again, the decay and intensity are empirically determined parameters. We set them, after testing, as follows:

$$Decay = 0.02, Intensity = 0.75.$$

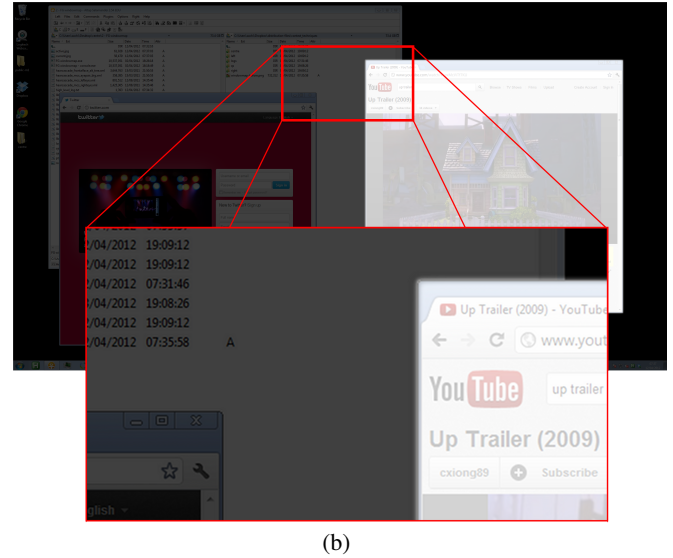
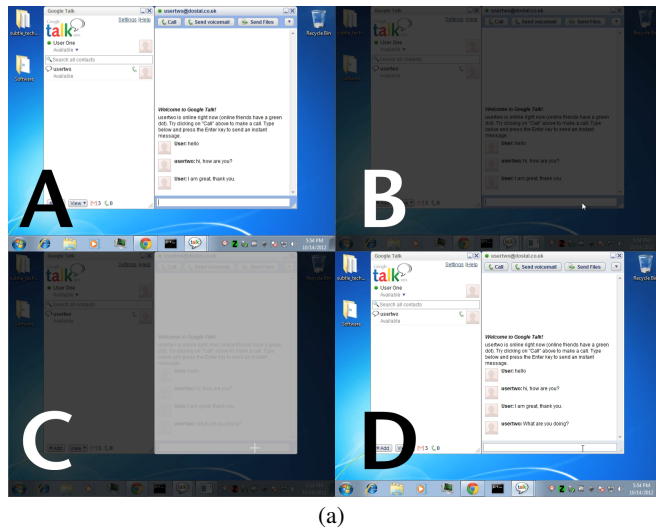


Figure 4. An illustration of WindowMap. It highlights visual change over time for individual application-windows. *A* shows the last frame before the user shifts visual focus away from this display. *B* shows a frame of the unattended display before any visual change. *C* illustrates how the technique visualises change (a new instant message in the chat window) as it happens by brightening the chat window. *D* shows the current display state after the technique dissolves. Figure 4b shows a detailed example of change visualised by brightening the application window, where the change occurred.

Aura

Aura visualises short-term display changes on an unattended display. Consider a user shifting their visual focus away from display ONE to Display TWO. Aura then first darkens display ONE to reduce distractions. Then Aura continually captures the last twenty frames of the unattended display at approximately one-second intervals. The change for a particular frame is then visualised as a thin rectangle around each window visible on the unattended display ONE. The brightness of the rectangle for each application window is proportional to how much the window in the frame changed in relation to the previous frame. Since Aura is tracking the changes for the last twenty frames, the visualisation ends up with twenty evenly spaced thin rectangles around each window. The closer a rectangle is to its window, the more recent the visualised change. Figure 5 and Video Figure 1.4 illustrate Aura. As the video figure illustrates, a window with many changes (such as a video player) results in cycles of bright rectangles, while a window with few changes (such as an instant messaging window) results in occasional bright rectangles.

DISPLAY-LEVEL VISUAL FOCUS DETECTOR

To enable the subtle visualisation techniques to be deployed to workstations, we developed a display-level visual focus detector using commodity hardware. Our system tracks users’ gaze using web cameras mounted on the displays. The algorithm is an enhanced version of an algorithm described in our previous work [9]. It uses a display-level gaze direction detector based on the Viola-Jones feature detection algorithm [22], a computer vision algorithm that is readily available in the open source OpenCV library. Our system uses three specific classifiers. These classifiers have been shown to achieve high accuracy [5]. The first classifier (EP1) detects

an eye-pair. The other two classifiers (LE and RE) detect the left and right eye respectively.

Using these classifiers we have designed a real-time algorithm that can detect eye-pairs with high accuracy and supports graceful degradation in situations when it is difficult to detect the user’s eye-pair. When the system cannot reliably identify the user’s left and right eyes, it attempts to infer an estimate based on an identified eye-pair area of interest. Algorithm 1 describes our gaze detection algorithm for a given image.

Evaluation

To ensure the visual focus detector was sufficiently accurate and robust to be deployed we carried out an evaluation with seven participants.

Method

The system was deployed to three dual-core iMac computers with 20-inch displays (1680×1050 pixel resolution). A Logitech C910 web camera was attached to each display. All computers were running the Windows 7 operating system. Each display was positioned at 76.5 cm from the user’s default sitting position. This distance was based on the results of a study of comfortable viewing distances for computer displays, which reported this distance as the most comfortable viewing distance on average [10].

We tested six angular differences (15°, 30°, 45°, 60°, 75° and 90°). Figure 6a illustrates how angular difference was defined in the evaluation. The six different angles were chosen to sample the whole visual field of the participants. We recruited seven university students as participants (6 male). Their ages ranged between 19 and 26. Their skin colour and hairstyle varied but none of them wore reading aids.

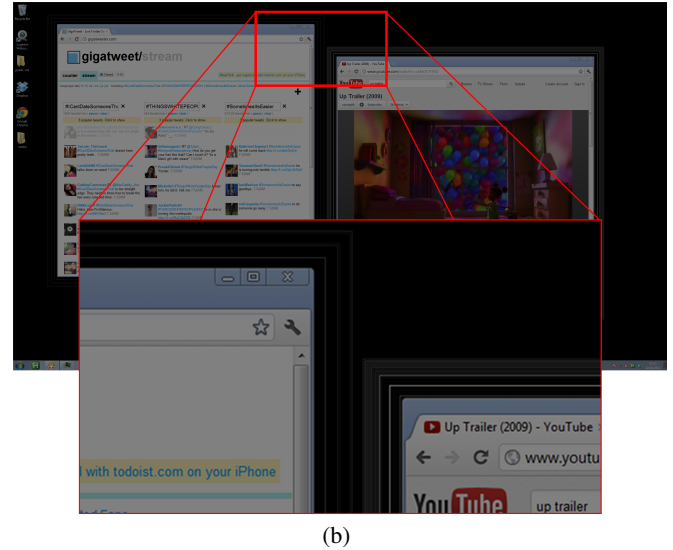
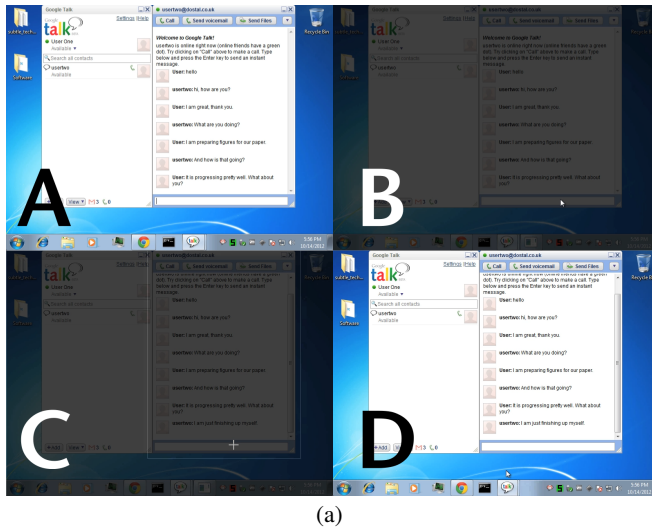


Figure 5. An illustration of Aura. It highlights short-term visual changes by projecting a rectangular aura around each window. *A* shows the last frame before the user shifts visual focus away from this display. *B* shows a frame of the unattended display before any visual change. *C* illustrates how the technique visualises the change (a new instant message in the chat window) as it happens as a rectangle around the chat window. *D* shows the current display state after the technique dissolves. Figure 5b shows a detailed example of change visualised as rectangles of varying brightness projected around the application window where the change occurred.

Algorithm 1 Algorithm used by the Visual Focus Detector

```

find all possible eye-pairs in a given camera image using
the EP1 classifier
for all detected eye-pairs do
    partition the rectangle containing the eye-pair horizon-
tally into two parts (left, right) and add 20% horizontal and
60% vertical padding to each of the two parts;
    scan the left and right sub-area of interest for potential
rectangles containing the left eye (using the LE classifier)
and the right eye (using the RE classifier), respectively, and
select the largest one in each case;
    if both sub-areas contain a detected eye then
        compute the pupil position for each of the eyes,
store their location and the containing rectangles for the
eyes;
    end if
end for
for all eye-pairs do
    if eye-pair has both eyes detected and the sizes of
the rectangles containing each eye match each other more
closely than current candidate then
        store this eye-pair as the best candidate;
    end if
end for
if there is a best candidate eye-pair then
    report a full detection, including the distance between
the pupils within the area of interest in pixels;
else[no best candidate is found]
    report a partial detection with the largest eye-pair found
as the most likely eye-pair;
end if

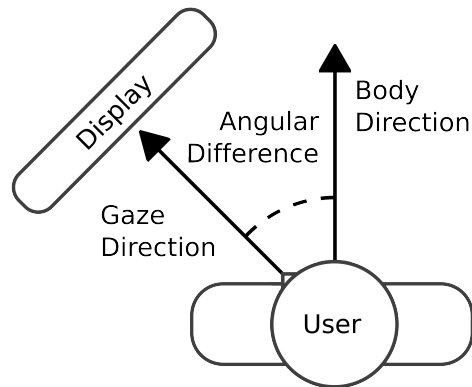
```

Each participant was asked to follow a sequence of spoken instructions directing them to look at a centre target (centre of a display) and side targets (centres of other displays) in turn. The participants were also instructed to keep their body facing the centre target and to only move their head and/or their eyes. As body position is not important during the actual detection, we wanted the participants to only move their head to exaggerate any possible effect imposed by the eye and head movement. They were asked to remain at the specified distance from the displays to maintain consistent viewing angles. For each examined angle, each participant spent approximately 25 seconds looking at the centre target and 20 seconds looking at each side target. This generated approximately seven minutes of video data per participant. Our algorithm then detected the participant's eyes for each video frame. Thereafter, the accuracy of our algorithm was assessed by visually comparing the algorithm's detection results against the ground truth in each image.

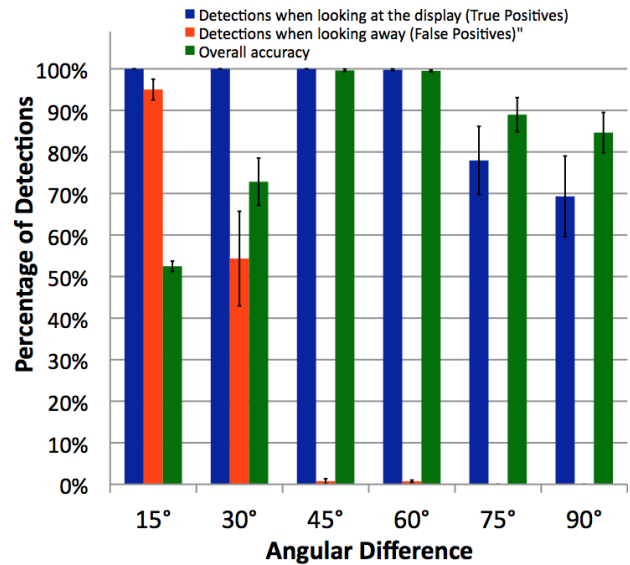
Results

Figure 6b shows the percentage of video frames with successful detections with our detection algorithm. The blue bars (first column) show the number of detections when the user is looking the centre of the display (true positives). The red bars (middle column) show the detections while the user is looking away from the display at a specific angle (false positives). The green bars (third column) show the overall accuracy of the algorithm.

For angular differences up to 45°, even though the algorithm performs perfectly in terms of true positives, the overall accuracy is significantly lowered due to the high number of false positives when the user is looking away. However, the amount of false positives steadily decreases as the angular dif-



(a) A diagram illustrating how angular difference was measured in the evaluation of the visual focus detector.



(b) The results of our study of horizontal detection angles. The error bars show standard error.

Figure 6. The experimental setup (a) and the results (b) for our visual focus detector.

ference increases, and at a 45° angular difference the overall accuracy is higher than 98%.

For extreme angles greater than 60°, detection accuracy decreases substantially. In these cases, the participants tended to not move their head much, but instead move their pupils all the way towards the corner of their eyes to see the display. This was expected as we specifically instructed participants to not move their body, which is what would naturally occur. This meant that for much of the time the participants' eyes appeared very different from how they typically appear when they look straight at a display.

While we used 20-inch displays in this evaluation, the results are display size independent as the targets were positioned in the centre of the displays. A 45° angular difference provided the highest performance. As a calibration point, two 27-inch displays, positioned at a comfortable viewing distance 76.5 cm away from the user bezel-to-bezel, have an angular difference close to 45° (see Figure 1). We used two 27-inch displays in our setup for the longitudinal study, which we describe next.

LONGITUDINAL CASE STUDY

It is difficult to evaluate the effects of our subtle visualisation techniques as their behavioural influences require long-term use in realistic contexts. As has been observed in the literature, controlled experiments are poor constructs for evaluating such techniques, and a naïve application of a quantitative usability study risks generating misleading results [11, 16].

Therefore, we instead conducted a formative five-day qualitative case study with a working professional. We recruited a single participant with no prior knowledge of the system or the techniques. This participant was 24 years old, familiar with software development and a regular user of multi-display

environments. The participant wore thin-rimmed glasses. Glasses do not have an effect on the display detector at short to medium distances. This effect only becomes noticeable at long distances when the eye-pair only occupies a small number of pixels in the camera image. Additionally, prior to the study, we extensively tested our display-level visual focus detector to ensure it would be able to accurately detect which display the participant attends to in his multi-display working environment.

The system was evaluated using a workstation configuration consisting of three computers, each with a high-resolution display. We used two 2010 iMac computers with 27-inch 2560×1440 pixel displays and a generic 2010 PC with a 24-inch 1920×1080 display. A Logitech C910 camera was attached to each display. Each camera provided a constant image stream of 1024×576 pixel images, which our system processed at 12 frames per second. All computers were running the Windows 7 operating system. Each display was positioned at 76.5 cm from the participant's preferred sitting position. However, the participant was instructed to sit comfortably and there was no restriction on their movement.

In this study, due to the size of the displays, the angular difference between the centres of the 27-inch displays was approximately 45° as measured from the 76.5 cm seating distance of the participant. This difference was maintained with the smaller 24-inch display for consistency. This setup represents a prototypical multi-display configuration, as a previous study has found that 45° angle between displays is a preferred setup among users [19].

Method

The participant was asked to use our multi-display setup for five days while performing his usual everyday work tasks.

The first day the participant was not exposed to a subtle visualisation technique. This approach was taken in order to obtain a baseline calibration point to compare the subtle visualisation techniques against. For the remaining four days, the four subtle visualisation techniques were deployed in the following order: WindowMap, FreezeFrame, PixMap and Aura. The participant was informed that some variant of an interaction technique that involves multiple displays would be used everyday but we did not inform the participant in advance of the particular technique that was used for a particular day. Further, the techniques were not demonstrated to the participant prior to the study.

Our system automatically logged the usage of specific displays, the time spent looking at each display, information about gaze switches between displays, the start and stop times for the techniques, and the method by which the technique was terminated (fade or user input).

The participant was also instructed to fill out two questionnaires each day. The first questionnaire was filled out immediately after the participant finished their work for the day and the second one the morning after. This meant we could ask the participant about both his immediate impression of a technique, and his impression of the technique the day after.

To complement the above methodology, we also used the Experience Sampling Method (ESM), which has been proposed as a methodology for evaluating pervasive and ubiquitous computing interfaces [6]. ESM enabled us to collect data of a higher granularity by presenting the participant with very brief questionnaires at approximately 30 minute intervals throughout the work day. The questionnaires asked about the participant's current task, his perceived frequency of display switching, and his familiarity with the technique. Finally, it included a comment area where the user could provide additional feedback. We collected between 9 and 15 samples per day, depending on the participant's physical presence and schedule. Additionally, one of the authors was present in the room for the entire duration of the experiment to administer the ESM questionnaires and to make visual observations.

Results

In total, we collected 39.25 hours of logged usage data (with a one second sampling rate), 9 questionnaires, and 61 ESM samples from our participant.

Display Utilisation

The participant spent a total of 31.66% of his time not attending to any of the displays due to various reasons such as meetings or paper-based work. Figure 7 shows the participant's display utilisation during the study as recorded in the system logs, adjusted to active-use time. Figure 7 makes it clear that the participant primarily attended to the centre display. This data suggests that the additional screens are mostly used for secondary or peripheral tasks. This is in line with previous research on the utilisation of multiple monitors [12]. In addition, based on our ESM samples and our visual observations, peripheral display tasks appear to be mostly related to social networking (Facebook, Twitter), instant messaging (MSN) or

media streaming (Spotify). The peripheral displays were also commonly used as a point of reference for information such as documentation, server logs or an application programming interface (API). The participant also referred to the displays on either side, as well as activities performed on them, as peripheral.

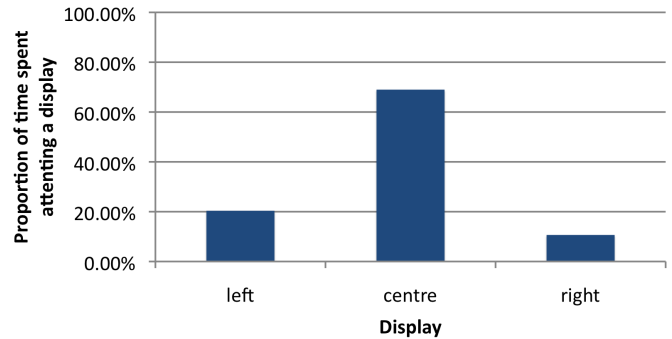


Figure 7. The distribution of display usage adjusted to active-use time.

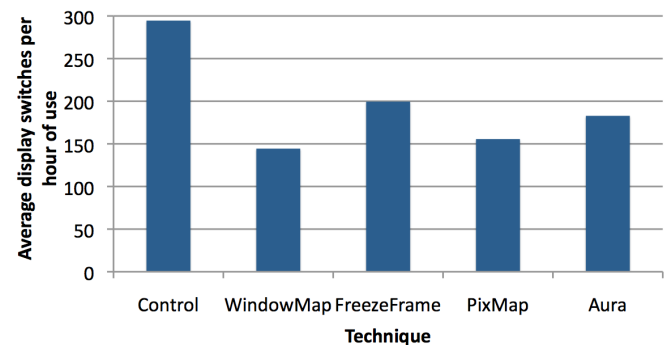


Figure 8. The average number of display switches per hour of use for each of the techniques.

Display Switching

Figure 8 plots the mean number of switches per hour of use for each technique. All four subtle visualisation techniques reduced the frequency of display switching by approximately a third or more compared to the baseline (which used no technique). FreezeFrame reduced the frequency of display switching by the smallest amount, only 32.5%. We conjecture this may be due to FreezeFrame's inability to dynamically visualise display change. This means that the only way for a user to check if there has been a change on the display is to switch visual focus to that display. Aura reduced the rate of display switching by 37.9%. We conjecture this may be due to the rectangular "aura" around the windows being too dim for the participant's peripheral vision to efficiently perceive it. Both PixMap and WindowMap reduced the rate of switching the most (47.2% and 51%, respectively). The similarity between the results of these techniques was expected due to their similar nature. We conjecture PixMap and WindowMap visualised change in such a way that the participant could use his peripheral vision to only react to changes that warranted his attention.

Our analysis of the subjective data suggests the participant had difficulty estimating his frequency of display switching.

According to the end-of-day questionnaires, the participant believed he was intensely switching displays during the day. His median rating on a seven-point Likert-type scale (“How often did you switch displays today? 1 = not at all, 7 = very often”) was 7, and his minimum rating was 5. However, the ESM samples overall median rating for the study was 4 on the same seven point scale. While this still indicates that the participant felt he switched displays at least moderately frequently, the ESM rating is markedly lower. The participant perceived the display switches that happened during the work day as being more frequent at the end of the work day compared to the ESM samples taken throughout the day. Further, while the participant reported a high rate of display switching at the end of each day, based on the ESM samples he was subjectively display switching less often when the subtle visualisation techniques were deployed.

The high number of display switches shown in Figure 8 might be related to distractions caused by applications on the peripheral displays. The participant repeatedly noted that he was often distracted by pop-up notifications and other animated content on the first day of the study (the day none of the subtle visualisation techniques were used). On the other hand, the participant did praise every tested technique for their distraction-reducing qualities, be it through desaturation, decrease in contrast/dimming, or by only highlighting areas of the screen where a change was happening. This subjective data corroborates the measured frequency of display switching collected from our system logs (Figure 8).

Tasks and Display Switching

As part of our ESM sampling, we collected information about the primary task the user was performing at the time the sample was taken. This gave us coarse-grained information about the temporal distribution of the tasks during the study. Each task (or application use) specified by the participant was classified into a broader task category (see Table 1 for a list of task categories). Most of the task categories are self-explanatory. *Academic Reading* encompasses reading academic papers in PDF format and searching for papers online. *Internet Browsing* includes all other activities that use the browser, most often reading blogs and using social media.

Since the ESM samples were administered at approximately 30 minute intervals, the system logs with display switching data were divided into 15 minute segments. The timestamp of each ESM sample was used as a centre of the period during which the specified task was performed. This allowed us to estimate the time spent on each of the specified primary tasks and relate it to the number of display switches. Aside from the coarse estimation, a further limitation of this data is that we have little information about the secondary tasks performed. From the ESM samples and the observations of the study administrator, we are aware that the secondary tasks most often included a combination of internet browsing and social media use, instant messaging, media playback and server management through a terminal application.

Due to the unrestricted use of the computers, not all tasks were performed every day. We have therefore aggregated the data for all techniques together. While this does not allow for

fine grained comparisons between techniques, it does allow us to see if there is a general trend towards distraction reduction for specific tasks (as all the subtle techniques should reduce distractions to some extent) and to compare different tasks with regards to the amount of display switching.

Task	Switches		Time [mins]	
	total	per min	active	total
Academic Reading	1035	2.43	425.42	540
Technical Reading	94	2.35	40.05	45
Graphical Editing	1207	3.33	362.72	450
Text Editing	583	2.45	237.52	315
Internet Browsing	200	3.07	65.08	105
Instant Messaging	308	5.13	60.08	105
Email	52	4.08	12.73	30
Other	133	3.19	41.72	45

Table 1. Display switches for specific tasks while the subtle visualisation techniques were running.

Table 1 shows the frequency of display switching for all the primary tasks performed during the four days the subtle techniques were used (as that is the larger, more diverse dataset). Apart from the total number of switches for a particular task, the table also shows the mean number of switches per minute of active computer use. In addition, the table shows the estimated total duration of each task over the four days, as well as the active computer usage time. The tasks with the highest mean number of switches per minute are the tasks which either tend to be multi-display (e.g. *Graphical Editing* with source information on one display and generated graphics on another) or tend to be of low importance, or glance-able (e.g. *Instant Messaging* and *Internet Browsing*). The high mean number of switches for email may be due to the short overall amount of time spent on the task and thus the result may not be very representative.

Task	Switches [per min]		Change
	Control	Techniques	
Graphical Editing	4.30	3.33	-22.62%
Text Editing	5.65	2.45	-56.56%
Instant Messaging	5.31	5.13	-3.40%

Table 2. The effect of subtle visualisation techniques on specific tasks.

Table 2 shows a comparison of three tasks that were performed both during the control day and the four days with subtle techniques. The table reveals the mean number of switches per minute for the two datasets. The subtle techniques decreased the number of display switches in all cases. In the case of *Instant Messaging* the change was small (3.4% decrease), which is likely due to the glance-able nature of instant messaging. However, for both of the editing tasks, the decrease is much more pronounced. The decrease for *Graphical Editing* is lower (22.62%) than for *Text Editing* (56.56%). This is probably because *Graphical Editing* is more expensive in terms of display space and thus is more likely to be used as a multi-display task, therefore requiring more frequent switches between displays. In either case, we can conclude that, especially for editing tasks, the subtle techniques successfully reduced the frequency of display switching.

Qualitative Analysis

In addition to investigating display utilisation and display switching, we interviewed the participant about the perceived usefulness of the techniques.

The first technique the participant used was WindowMap. Our observations indicate that the technique helped reduce distractions and allowed the participant to better focus on his primary task. For example, the participant stated that “*Dimming peripheral screens makes it easier to focus on [the main screen].*”

However, the timing design of WindowMap confused the participant. WindowMap was set up so the visualised changes would decay after approximately 20 to 30 seconds if no other changes occurred. This seems to have led to a situation, where the participant was noticing a change but was unsure about the timing of when the change occurred and whether or not it was continuous. However, we noticed a possible reduction in cognitive load when the participant processed events from the peripheral displays. For example, the participant stated that “*while waiting for a window to finish processing something, I was able to note the state change in my peripheral vision, which was more useful than randomly glancing over to check if it was done.*” The participant ranked this subtle visualisation technique third in terms of perceived usefulness.

The second technique the participant used was FreezeFrame. FreezeFrame is a technique for re-establishing prior context before and after a display switch. FreezeFrame requires full visual focus and does not rely on users being able to notice changes in their peripheral vision. Our qualitative data indicates that this behavioural design of FreezeFrame helped the participant monitor longer-term changes more effectively: “*I find it very useful for more passive applications such as twitter. When glancing over, I have an immediate sense of what tweets are new, without being constantly distracted.*” However, the fact that it required the participant’s full visual focus meant that checking peripheral, yet time-sensitive applications, such as an instant messaging conversations created more of a distraction: “*I’ve found it useful to be able to glance at contact lists to see who’s come online recently, because the change in the list is so easy to perceive, without distracting me while working.*” Finally, while the participant felt the time-to-fade was too long, it was still considered the second most useful technique.

The third technique the participant used was PixMap. As previously noted, PixMap is similar to WindowMap in that it visualises changes over time. However, while WindowMap visualises display changes over time on the application-window level, PixMap visualises display changes over time on the pixel level. The participant perceived PixMap as more useful than WindowMap because it provided him with more fine-grained information, while also being more subtle. The participant reported that PixMap was particularly useful for applications that only resulted in smaller display changes, such as list updates: “*The ‘trail’ effect is particularly useful for gauging how lists have changed, such as on twitter - giving a glanceable way of telling how many new tweets there are.*”

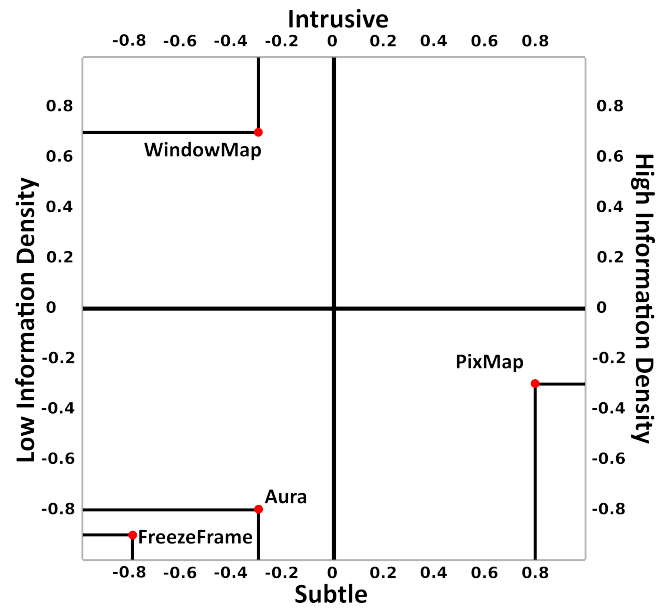


Figure 9. Mapping of all the techniques based on their perceived subtlety/intrusiveness (y -axis) and information density (x -axis).

The participant ranked PixMap as the most useful of all the subtle visualisation techniques.

The last subtle visualisation technique the participant tested was Aura. The participant felt it was useful for reducing distractions: “*Not distracting - keeps me focused!*”. Aura was perceived as very subtle, sometimes too subtle. In part this was because the visibility of the visual “pulses” that visualised the amount of change in the window at each of the past time steps depended on the background colour of the application. Since high activity was visualised as a white colour, this meant that when the background was white, it became difficult to perceive the change: “*I’m aware of change in some apps more than others. For example, one screen has two apps with dark colour schemes so the ‘pulse’ effect is more pronounced, whereas with apps with white backgrounds, I often miss any subtle change.*” The participant ranked Aura as the least useful subtle visualisation technique.

Subtlety and Information Density

Figure 9 shows the participant’s subjective mapping of the techniques in terms of their subtlety or intrusiveness and information density. Information density is a measure of the compactness of a particular technique in terms of the amount of useful information the technique displays. We noted that the ordering on the subtlety axis closely corresponds to the ranking of perceived subtlety reported by the participant in the daily questionnaires.

There was a separate question in the questionnaires about perceived contribution to increased productivity. Its ranking was exactly the same as for perceived usefulness. Since periphery-enhancing information is seen as one of the indicators of calm technologies [23], we expected the perceived usefulness of a technique to correspond with its perceived information density. However, this was not the case as the

ranking of perceived usefulness was No Technique < Aura < WindowMap < FreezeFrame < PixMap, while for information density it was No Technique < FreezeFrame < Aura = WindowMap < PixMap. This means that while FreezeFrame was ranked low on information density, it was perceived as very useful.

LIMITATIONS AND FUTURE WORK

One limitation with our work is the way the system detects users' visual focus on a particular display. Our computer vision methods are dependent on inter-device angular difference. Since our system relies on a minimum angular difference between displays for accurate visual focus detection, users' multi-display environments may need to be modified to accommodate this. This can be done by either changing the size of displays, or their distance from the user.

Another limitation is that we rely on a five-day case study when we evaluate the implications and limitations of the four subtle visualisation techniques. Our choice to rely on a qualitative in-depth case study involving a single participant is unusual, and perhaps somewhat controversial. Studying a single participant limits our ability to make statistical inferences about the performance of subtle visualisation techniques for the general user population. However, in terms of the amount of data we collected, our 39.25 hours of data is approximately 2–3 times as much as what we would have collected if we carried out a traditional one-hour within-subjects controlled experiment with 16 participants. More importantly, it is very difficult to design a meaningful controlled experiment to evaluate the efficacy of subtle visualisation techniques. Unlike many other intelligent user interfaces, such as adaptive user interfaces and pointing amplification techniques, subtle visualisation techniques lack meaningful measures that can be collected within a relatively short amount of time. Further, lab-based studies de-contextualise the interaction techniques. For many intelligent interaction interfaces, this may not be a major issue. However, our subtle visualisation techniques are designed to be non-intrusive in users' regular working environments. There is therefore a substantial risk that de-contextualising the working environment in a laboratory study would have provided us with misleading results. In contrast, ESM enabled us to get a nuanced understanding of how our working professional perceived subtle visualisation techniques in a multi-display environment while carrying out his regular work.

Our work suggests that while our evaluation method works, it is also expensive. Therefore, a fruitful avenue for future work would be to develop alternative, more cost-effective, ways of evaluating subtle visualisation and interaction techniques. One possible direction is the design and evaluation model suggested by McCrickard et al. [15]. Although this model concentrates on notification systems, it could perhaps be broadened and adapted for use as a benchmark for evaluating subtle interaction techniques.

CONCLUSIONS

This paper explored subtle gaze-dependent techniques for visualising display changes in multi-display environments. We

presented four subtle gaze-dependent techniques for visualising changes in unattended displays: FreezeFrame, PixMap, WindowMap and Aura.

To enable the visualisation techniques to be directly deployed to workstations, we also presented a system that automatically identifies the user's eyes using computer vision and a set of web cameras mounted on the displays. Our evaluation confirmed this system can accurately detect which display the user is attending to.

We studied the efficacy of the visualisation techniques in a longitudinal case study with a working professional. This participant used our system eight hours per day for five consecutive days. The results of the study showed that the techniques were successful in highlighting visual change, and in reducing distractions and the frequency of display switches for the participant. Further, three of the four techniques did so in a non-intrusive and subtle manner.

We also noted some challenges when evaluating subtle intelligent interaction techniques that require deployment in actual working contexts in order to generate meaningful data. In our study we used a combination of questionnaires, ESM sampling, observations and automatic logging. Our methodology proved successful for our case study, but it is also prohibitively expensive for a large-scale study. Our work points to potential avenues for overcoming these limitations through the use of an alternative evaluation methodology.

Our case study suggests that subtle visualisation techniques can indeed positively change the display attendance behaviour of a working professional who used these techniques as part of his regular work activities. However, many opportunities still exist for taking the technology further, both in terms of further refining the techniques, and in terms of identifying cost-effective means of evaluating the efficacy of such techniques. The subtle display of visual change information from unattended displays can alter our interactions and expectations of interfaces. Out of sight no longer means out of mind, when inattention is no longer an inherent cost but a new research and development opportunity.

ACKNOWLEDGEMENTS

This work was supported by the Engineering and Physical Sciences Research Council (grant number EP/H027408/1) and the Scottish Informatics and Computer Science Alliance.

REFERENCES

1. Ashdown, M., and Sato, Y. Attentive Interfaces for Multiple Monitors. In *CHI 2005 Workshop on Distributed Display Environments*, 4–5.
2. Baudisch, P., Good, N., Bellotti, V., and Schraedley, P. Keeping Things in Context: A Comparative Evaluation of Focus Plus Context Screens, Overviews, and Zooming. In *Proc. CHI 2002*, 259–266.
3. Beck, M. R., Levin, D. T., and Angelone, B. Change Blindness Blindness: Beliefs about the Roles of Intention and Scene Complexity in Change Detection. *Consciousness and cognition* 16, 1 (2007), 31–51.

4. Bi, X., and Balakrishnan, R. Comparing Usage of a Large High-Resolution Display to Single or Dual Desktop Displays for Daily Work. In *Proc. CHI 2009*, 1005–1014.
5. Castrillón-Santana, M., Déniz-Suárez, O., Antón-Canalís, L., and Navarro-Lorenzo, J. Face and Facial Feature Detection Evaluation: Performance Evaluation of Public Domain Haar Detectors for Face and Facial Feature Detection. In *Proc. VISAPP 2008*, 167–172.
6. Consolvo, S., and Walker, M. Using the Experience Sampling Method to Evaluate Ubicomp Applications. *IEEE Pervasive Computing* 2, 2 (2003), 24–31.
7. Czerwinski, M., Smith, G., Regan, T., Meyers, B., and Starkweather, G. Toward Characterizing the Productivity Benefits of Very Large Displays. In *Proc. INTERACT 2003*, 9–16.
8. DiVita, J., Obermayer, R., Nugent, W., and Linville, J. M. Verification of the Change Blindness Phenomenon While Managing Critical Events on a Combat Information Display. *Human Factors* 46, 2 (2004), 205–218.
9. Dostal, J., Kristensson, P. O., and Quigley, A. Estimating and Using Absolute and Relative Viewing Distance in Interactive Systems. *Pervasive and Mobile Computing* (2012). doi:10.1016/j.pmcj.2012.06.009
10. Grandjean, E., Hünting, W., and Nishiyama, K. Preferred VDT Workstation Settings, Body Posture and Physical Impairments. *Journal of Human Ergology* 15, 2 (1984), 99–104.
11. Greenberg, S., and Buxton, B. Usability Evaluation Considered Harmful (Some of the Time). In *Proc. CHI 2008*, 111–120.
12. Grudin, J. Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use. In *Proc. CHI 2001*, 458–465.
13. Intille, S. S. Change Blind Information Display for Ubiquitous Computing Environments. In *Proc. UbiComp 2002*, 1–16.
14. Kern, D., Keynes, M., and Schmidt, A. Gazemarks - Gaze-Based Visual Placeholders to Ease Attention Switching. In *Proc. CHI 2010*, 2093–2102.
15. McCrickard, D. S., Chewar, C. M., Somervell, J. P., and Ndiwalana, A. A Model for Notification Systems Evaluation - Assessing User Goals for Multitasking Activity. *ACM Transactions on Computer-Human Interaction* 10, 4 (2003), 312–338.
16. Olsen, D. R. J. Evaluating User Interface Systems Research. In *Proc. UIST 2007*, 251–258.
17. Shneiderman, B., and Benderson, B. B. Maintaining Concentration to Achieve Task Completion. In *Proc. DUX 2005*, 2–7.
18. Simmons, T. What's the Optimum Computer Display Size? *Ergonomics in Design* 9, 4 (2001), 19–25.
19. Su, R. E., and Bailey, B. P. Put Them Where? Towards Guidelines for Positioning Large Displays in Interactive Workspaces. In *Proc. INTERACT 2005*, 337–349.
20. Vertegaal, R., Dickie, C., Sohn, C., and Flickner, M. Designing Attentive Cell Phones Using Wearable EyeContact Sensors. In *CHI Extended Abstracts 2002*, 646–647.
21. Vertegaal, R., Shell, J. S., Chen, D., and Mamuji, A. Designing for Augmented Attention: Towards a Framework for Attentive User Interfaces. *Computers in Human Behavior* 22, 4 (2006), 771–789.
22. Viola, P., and Jones, M. J. Robust Real-Time Face Detection. *International Journal of Computer Vision* 57, 2 (2004), 137–154.
23. Weiser, M., and Brown, J. S. Designing Calm Technology. *World Wide Web Internet And Web Information Systems* 1, 1 (1995), 1–5.