Visual momentum redux

Kevin B. Bennett*, John M. Flach

Psychology Department, 335 Fawcett Hall, Wright State University, Dayton, OH 45435, United States

Received 5 August 2010; received in revised form 20 January 2012; accepted 22 January 2012
Communicated by D. Boehm-Davis
Available online 23 February 2012

Abstract

Over 25 years ago Woods (1984) introduced the concept of visual momentum: the extent to which an interface supports a practitioner in transitioning between various information-seeking activities that are required for understanding and exploring work domains. Increasing visual momentum requires the consideration of a range of “cognitive couplings” that span all levels of the interface: between multiple screens, within individual screens, and within a display on a screen. Although the concept has been well received, we believe that its potential to improve the quality of human computer interaction may be under-appreciated. Our purpose in this review is to provide a better understanding of visual momentum: to provide concrete and diverse examples of its successful application, to review empirical findings, to refine and expand the original design techniques that were proposed, and to integrate diverse terms that appear across different research communities.

© 2012 Published by Elsevier Ltd.

Keywords: Visual momentum; Cognitive systems engineering; Long shot; Focus+context; Overview+detail; Fish eye views; Center surround; Landmarks; Ecological interface design

1. Introduction

An area that has received little attention in the design of computer-based display systems is how the user integrates data across successive displays (cf. Badre, 1982) … Failure to consider the requirements for effective across-display user information processing can produce devastating effects on user performance … The “getting lost” and “keyhole” phenomena are not inevitable consequences of using computer-based displays; neither do they represent human limitations (for example, short-term memory) that must be compensated for through memory aids or walls of VDUs. Across-display processing difficulties are the result of a failure to consider man and computer together as a cognitive system … (Woods, 1984, pp. 229–230).

Today’s work domains are complicated, often involving large, complex workspaces and databases. This is true for traditional work domains (e.g., process control, military command and control) and work domains as more broadly defined (e.g., consumer electronics). Interaction in these work settings has become increasingly “computer mediated” over time: the advent of computational technologies has resulted in the computer interface becoming the practitioner’s primary window on the world of work. Since the physical size of viewing surfaces in the interface is very limited, this window is often an extremely small one.

This combination (i.e., complex work domains and inherently limited “display real estate”) has had a fairly direct impact on the ability of practitioners to conduct work effectively. All of the information required for effective control cannot possibly be displayed in parallel; the practitioner must selectively view glimpses serially, over time, through the limited “keyhole” provided by the small display surface (i.e., the “keyhole” effect, Woods, 1984). Thus, information must often be remembered and mentally integrated across successive glimpses, an activity that practitioners are not particularly well equipped to handle.

A direct consequence is that practitioners must navigate these complicated workspaces and databases to find
information that is relevant for the task at hand. To navigate successfully, the practitioners must know their current location within the workspace, the space of possible locations that might be navigated to, and a navigational route that will get them there. Unfortunately, information systems rarely provide interface resources that support these needs and the end result is often what Woods (1984) referred to as the “getting lost” phenomenon: “… users often do not know where they are, where they came from, or where they have to go next” (Ziefle and Bay, 2006, p. 395).

We recently had the opportunity to observe an example at first hand. In 2007 we participated in exercises for modern air combat operations. The work setting was large (auditorium-sized), collaborative (nearly 100 people) and it featured cutting-edge information technology (e.g., display walls, chat rooms, etc.). Our observations include the following. Even simple operations often required practitioners to navigate across multiple display windows and to integrate information (sometimes presented in different formats). Multiple overlapping windows occluded, and sometimes even locked out, critical information. Practitioners were unable to perform reasonably simple (and critical) tasks in some instances because of the lack interface support. “Work-arounds” (e.g., jotting down critical information using pen and paper as a memory aid) were common. Ironically, despite cutting edge technology (e.g., display walls), the most common complaint was the need for more screen real estate.

Other researchers have observed similar problems across a variety of work settings and information technologies. Examples include studies of cell phones (Zieff and Bay, 2006), database queries (Vicente ET AL., 1987), medical systems (Cook and Woods, 1996; Obradovich and Woods, 1996), process control (Easter, 1991; Elm and Woods, 1985), aviation (e.g., Aretz, 1991), spreadsheets (Watts-Perotti and Woods, 1999) and hypertext documents (McDonald and Stevenson, 1996).

Over a quarter of a century ago Woods (1984) explicitly defined the fundamental problems and issues (i.e., complex worlds, limited display real estate, the keyhole effect and the getting lost phenomenon) and the associated challenges for system design. Despite significant advances in interface technologies (or perhaps due to it), problems like the keyhole effect and the getting lost phenomenon have become even more commonplace in the interim.

A primary factor that appears to have contributed to the lack of progress is what we will refer to as the “myth of information technology”: an unwarranted confidence that the use of computational technologies alone will be sufficient to address these challenges. For example, a belief that simply increasing the size of the window on the world (e.g., data walls) will alleviate problems like the keyhole effect. Or, a belief that artificial intelligence can be applied to automate data management and presentation (e.g., intelligent interfaces) or even task performance (e.g., expert systems) to alleviate problems like the getting lost phenomenon.

1.1. Visual momentum (VM)

There is an alternative to devising and throwing new forms of technology at these problems. Woods (1984, p. 231) proposed the concept of “visual momentum,” defining it as “… a measure of the user’s ability to extract and integrate information across displays…” In this approach existing interface technologies, both powerful and mature, are used wisely with the explicit goal of increasing the degree of “cognitive coupling” between practitioners and the work domain. Note that the term was originally coined by Hochberg (1986) (Hochberg and Brooks, 1978) who studied the perception of cinematic film.

The need to support visual momentum is most obvious at the highest levels of structure in the interface (i.e., when an interface has multiple screens). Interaction at this level, the “workspace” level, involves transitions between successive computer screens or pages: one collection of displays on a screen is replaced with another. The degree of visual momentum in an interface is determined by the “smoothness” of these segues. Do the practitioners know where they are, where they have been, and where they might want to go? Does the interface prepare the practitioners (i.e., set up an appropriate set of expectations) for this transition by perceptually and/or cognitively orienting them to the particulars of their new destination?

There is also a need to support visual momentum at lower levels of structure in the interface, including transitions within a screen or page (i.e., between collections of displays or windows—the “view” level) or within a display (i.e., between the visual elements of a display itself—the “form” level). This need is perhaps less obvious due to our phenomenological experience of a stable and unbounded visual world (e.g., Gibson, 1966), which here refers to the subjective impression of parallel data presentation that one experiences when looking at a computer screen or display. However, this impression does not match the anatomical and physiological facts (e.g., blind spot, eye movements, field of view). In the end it is clear that the visual system can only obtain information about the world by scanning it in a serial fashion (and therefore, that visual momentum techniques are needed to support transitions at this level).

1.1.1. Techniques to increase visual momentum

Woods (1984) originally proposed and then refined (Watts-Perotti and Woods, 1999; Woods and Watts, 1997) a set of design strategies or techniques to increase visual momentum. Similar techniques have been proposed and evaluated by researchers in the human computer interaction (HCI) and computer science communities, although these researchers have chosen a different set of descriptive labels. Fig. 1 (adapted from Woods, 1984) organizes these techniques in terms of their potential to increase visual momentum and integrates interdisciplinary terminology.

The purpose of this article is to reexamine the concept of visual momentum and the progress that has been made in
its application. One goal is pedagogical in nature: to provide additional descriptions and concrete examples of these techniques and how they have been applied in a variety of different work domains. A second goal is practical in nature: to review instances where these techniques have been implemented successfully and to document the benefits to overall human-machine performance that have been obtained. The article will be structured around the techniques in Fig. 1, starting at the left and working towards the right.

2. Fixed format data replacement

A simple (yet powerful) example of the “fixed format, data replacement” technique is spatial dedication. A natural consequence of interfaces designed with older-generation, hardwired technology was that controls and displays were spatially dedicated. In contrast, with computerized interfaces this becomes a design goal: care must be taken to place interface objects (e.g., menus, commands, categories of information, display frames, labels, etc.) in specific and consistent spatial locations (both within and across views). Spatial dedication increases visual momentum via a form of cognitive orientation. Practitioners will learn where specific objects and information are located as they gain experience with an interface; required transitions are facilitated because accessing information and executing control input are more efficient.

Spatial dedication appears to be a compelling and general requirement for effective interface design. For example, Woods and Watts (1997, p. 642) have observed practitioners “... throwing away flexibility ... to create their own spatially dedicated set of views of the underlying process or device.”

2.1. Example: RAPTOR interface

The “signature” version of the fixed format data replacement technique allows practitioners to selectively change the informational content of a display without changing the viewing context that it resides within (e.g., surrounding information; general format of the display including axes, units, labels). An example will be drawn from our work (e.g., Bennett et al., 2008; Hall et al., in press; Talcott et al., 2007) on an ecological interface (RAPTOR, Representation Aiding Portrayal of Tactical Operations Resources) for military command and control. Fig. 2 illustrates a coordinated set of displays used to represent the combat resources (e.g., ammunition, fuel, personnel, tanks, and armored personnel carriers) of an Army battalion. Note the hierarchically nested organizational structure (i.e., one battalion, four companies, 12 platoons, and 48 vehicles) of the battalion that is represented in the tree structure of the aggregation control/display at the top of Fig. 2a.

The fixed format data replacement technique is used in the primary and secondary display slots (see the labels in Fig. 2a). The number and content of the displays that appear in these two slots are variable. The primary display slot represents the combat resources of a superordinate unit (e.g., the battalion in Fig. 2a); the secondary display slot represents the combat resources of a subordinate unit (e.g., the companies in Fig. 2a). Rolling over (or pointing at and clicking on) a node in the tree structure replaces the displays in both slots with the appropriate displays for that particular node. Fig. 2b (company level) and 2c (platoon level) illustrates changes in the content of the slots for views at lower levels in the organizational structure (note that one more level is possible, but not illustrated—vehicles and the soldiers they contain).

This is a prototypical example of the fixed format data replacement technique: the entire database is too complex to be viewed in parallel; a limited amount of display real...
estate is used to provide detailed, selective, and serial glances into the database. The primary and secondary slots are located in fixed areas of the interface (i.e., spatial dedication). All representations appearing in these slots utilize variations of the same general display format (e.g., horizontal bar graphs, color-coding conventions, boundaries), modified as needed to meet data requirements at the different levels (fixed format data replacement). Thus, with a single control input the practitioner can choose from 65 different perspectives on the database, allowing access to information that ranges from summarized combat resources for the battalion all the way down to the status of hundreds of individual soldiers.

2.2. Summary

In general, the fixed format data replacement technique increases visual momentum by supporting alternative,
manageable glances into a larger workspace. This allows the practitioner to selectively focus attention on relevant subsets of information (e.g., combat resources at various echelon levels). Visual momentum is increased in several ways. First, the fixed formats (e.g., primary and secondary display slots) provide spatially dedicated locations that frame the presentation of information. This limits the cognitive effort required for search (i.e., consistent mapping; automatic processing). Second, the new informational content is made available in parallel within the context of the original view. The practitioner does not need to navigate to an alternative view, remember the new data, and then return to consider what those data mean within the broader context (e.g., combat resources in light of extensive tactical information presented in other areas of the RAPTOR interface). Third, the transition between old and new information is facilitated by a common visual context between views (e.g., the same general display format that is customized for different organizational levels in RAPTOR). This reduces the amount of cognitive effort that is required for reorientation; the focus can be on changes in the data instead.

We are not aware of any empirical research that has addressed the fixed format data replacement technique directly. The technique is likely to have contributed heavily to significant performance advantages found for the RAPTOR interface (Hall et al., in press), but it was not an isolated manipulation. There is some empirical evidence supporting a very closely related presentation technique. Tufte’s (1990) “small multiples” information visualization technique embodies exactly the same principle (i.e., changes in information content displayed in the same viewing context). Several empirical evaluations (Archambault et al., 2011; Javed et al., 2010; Robertson et al., 2008) have found significant performance advantages for this technique. The only difference between the small multiples and fixed format data replacement techniques is that the former are presented in parallel, while the latter are presented serially.

3. Long shot (overview + detail); zoom + pan

In this section we describe two complementary techniques that increase visual momentum by changing the resolution or focus of the information that is being considered. The long shot design technique (Watts-Perotti and Woods, 1999; Woods, 1984; Woods and Watts, 1997) comprises two distinct components. One component provides an abstract overview of the global “space of possibilities” of a workspace. The second component provides, in parallel, a more detailed view of a specific portion of the overall workspace. The HCI/computer science communities have used the term “overview + detail” (Beard and Walker, 1990; Cockburn et al., 2008; Plaisant et al., 1995) to describe the technique, explicitly referring to both essential components.

The second technique discussed in this section is zoom + pan, which also has two distinct components. One component allows the practitioner to change the resolution of the current glance into a workspace by zooming in (e.g., magnification) or zooming out (e.g., demagnification). The second component allows the practitioner to shift the focus of the current glance into a workspace (i.e., pan). This is a powerful technique that is a routine feature of many applications.

3.1. Examples: RAPTOR and Adobe® Photoshop®

Fig. 2a illustrates the long shot technique as implemented in RAPTOR. The overview component is embodied in the aggregation control/display. This overview captures the hierarchically nested, whole-parts organization of the battalion, as described previously. Each node is color-coded (using the Army convention of green, amber, red, and black to represent successively fewer resources) to provide a categorical summary of the combat-readiness of each unit or vehicle. The node of the currently selected view (e.g., the battalion level in Fig. 2a) is differentiated from all other nodes by distinct perceptual encodings (a brighter color and a bolder outline). The detail component (i.e., graphical displays in the primary and secondary slots) and control mechanisms for changing focus were described in the previous section.

Fig. 3a illustrates the long shot technique as implemented in Adobe Photoshop. The overview component is located in the upper right portion of the screen. This overview (shown in enlarged proportions on the right side of the figure) presents a low-resolution view of the entire graphics file that is being edited. The detail component (the majority of the application screen) contains the portion of the overall graphics file that is currently being worked on. The boundary rectangle in the long shot display specifies the current boundaries of this detailed view.

Fig. 3b and c illustrates how the zoom + pan technique has been integrated with the long shot technique. In Fig. 3b the practitioner has changed the resolution of information in the detail view by zooming in. Note that the size and location of the boundary rectangle in the long shot display has changed to define the currently viewed portion of the graphics file. In Fig. 3c the practitioner changed the focus in the detail view by panning to the right in the long shot. In this case the pan has been accomplished via direct manipulation (i.e., the practitioner points, clicks, and drags the rectangle to another physical location within the long shot display).

Woods and Watts (1997) describe three primary functions that a well-designed long shot (or overview + detail) will provide: orienting, movement, and status summary. The orienting function refers to the provision of interface resources that illustrate the space of possibilities and a practitioner’s current location within that space (e.g., the long shot displays in Figs. 2 and 3). The movement function refers to the capability to switch the focus...
and/or the resolution of a particular glance into the overall workspace (e.g., the manipulations of the long shot displays).

The status summary function refers to the provision of information that “… allows users, in a mentally economical way, to step back and assess their overall situation with respect to the underlying process, device, or activity they are engaged in” (Woods and Watts, 1997, p. 637). To be successful, status summary information must meet several criteria: it must be distilled, abstracted, dynamic (with regard to change), relevant, and economical. This function is fulfilled in RAPTOR through the visual appearance of the tree nodes in the long shot display. These nodes specify the organizational structure of the
battalion (i.e., relevant). The color-coding of each node provides a categorical summary (i.e., distilled, abstracted) of associated combat resources at each level. The color-coding is dynamically updated (i.e., a salient cue illustrating change in status). Thus, at a glance, the practitioner can see interrelated summaries of status throughout the entire organizational structure (i.e., economical).

3.2. Summary

The long shot display technique “... provides an overview of the display structure as well as summary status data. It is a map of the relationships among data that can be seen in more detailed displays and acts to funnel the viewer’s attention to the ‘important’ details” (Woods, 1984, p. 236). The long shot technique increases visual momentum by supporting transitions between alternative glances into the overall workspace. The “getting lost” phenomenon is avoided by providing an overview of the overall layout of the workspace and by highlighting the particular “glance” that is currently being displayed in detail (e.g., “you are here”). The “keyhole” effect is avoided by providing integrated sets of selective and detailed glances that correspond to the various perspectives that might need to be considered (thereby meeting the dual challenge of complexity and limited display real estate). Visual momentum is increased because practitioners know where they are currently located, where they might go to in the future, and how intuitive control mechanisms can be used to get them there easily.

The long shot (overview+detail) technique has been analyzed, implemented, and evaluated extensively. Hornbaek et al. (2002) and Cockburn et al. (2008) provide summaries of the empirical evaluations in the HCI literature. In the human factors literature, the long shot technique is often embedded within large-scale interface design and evaluation efforts (e.g., the RAPTOR example and its evaluation, Hall et al., in press). However, several empirical studies have evaluated the long shot technique in a reasonably direct fashion (Billingsley, 1982; Errington et al., 2005; Roth et al., 1998; Tharanathan et al., 2010). In general, well-designed long shot displays have been found to enhance performance in a variety of settings.

Empirical evaluations of the zoom+pan technique and its variations (zoom+tilt+pan; integrated zoom+pan) have produced results that are somewhat mixed (e.g., Hornbaek et al., 2002). Literature reviews (e.g., Bederson, 2011; Cockburn et al., 2008) have found practitioner acceptance (practitioners generally like the technique), some performance benefits (particularly when well-implemented) and some performance trade-offs (e.g., disorientation after zooming). Recent work (Morison et al., 2009) has applied VM to go beyond single video camera zoom+tilt+pan techniques to aid in coordinating multiple perspectives, particularly of feeds from multiple cameras (surveillance or robots) when practitioners are removed from the scene of interest.

4. Perceptual landmarks

A third design strategy for increasing visual momentum is to provide perceptual landmarks in the interface. Siegel and White (1975) describe “landmark” knowledge in the following manner (p. 23): “Landmarks are unique configurations of perceptual events (patterns). They identify a specific geographical location. The intersection of Broadway and 42nd Street is as much a landmark as the Prudential Center in Boston ... These landmarks are the strategic foci to and from which the person moves or travels ... We are going to the park. We are coming from home.” This concept is directly applicable to display and interface design. Perceptual landmarks in the interface are representations that (1) occupy dedicated spatial locations and (2) provide cues with regard to meaningful aspects of the underlying work domain to which one can navigate.

4.1. Example: the BookHouse interface

The BookHouse system (Pejtersen, 1980, 1988, 1992) will be used to illustrate landmarks in the interface. This system equates the process of finding a book of fiction in a public library to the process of navigating through a virtual library. Fig. 4a provides a representative navigational sequence. The practitioner initiates a search by entering the virtual library (pointing at and clicking on the entrance). The entry hallway has three adjoining rooms (Fig. 4b) representing alternative sections of library holdings (e.g., children’s section), which can be searched. The practitioner narrows the search by navigating to one of these rooms (Fig. 4c). This room contains patrons, involved in activities, which represent alternative search strategies (e.g., search by analogy) that can be executed. The practitioner chooses a search strategy (clicks on a patron) and enters another room (Fig. 4d) containing objects (e.g., the globe, the clock, the eyeglass icons) that represent the various dimensions of search (e.g., geographical setting, time period, font size) that can be specified.

Thus, many of the spatial metaphors in the BookHouse interface (e.g., virtual library, rooms, patrons, objects) qualify as landmarks (i.e., “configurations of perceptual events”) that point to locations in a virtual workspace and associated functionality. The landmark technique is perhaps most clearly evident in the “navigation bar” located at the top portion of the screen. A metaphor is placed in the navigation bar each time the practitioner leaves an area of the virtual library (compare the contents of the navigational bar in Fig. 4a–d). The physical appearance of these metaphors mimics the physical appearance of the area being left (i.e., it provides a small-scale replica of the corresponding room, hallway, or entrance). These “replica” metaphors also serve as controls that can be pointed at and clicked on to produce navigation back to that location. Thus, the replica metaphors in the navigation bar provide perceptual landmarks that specify locations in the virtual workspace that can be navigated to so that an ongoing search can be modified.
The BookHouse interface uses a global metaphor that is explicitly tied to a real world setting. It should be emphasized that this is not necessary for implementation of the landmark technique. Consider the interfaces of Apple’s iPhone® and iPad® devices. A “home page” organizes application metaphors by placing them in a matrix. Multiple home pages can be configured (and navigated to). Tapping an application metaphor in a matrix produces navigation to the application. Each application, in turn, has metaphors for objects or application modes that can be also be tapped (with ensuing navigation to associated functionality). Note that each level of spatial organization provided by the real world setting used in the BookHouse interface (i.e., library, rooms, objects in the room) has a direct parallel in the spatial organization of the iPhone and iPad interfaces (i.e., multiple home pages, individual home pages, application objects or modes). Thus, each metaphor in the iPhone and iPad interface (application, mode, or object) essentially provides a perceptual landmark to associated functionality in the workspace.

4.2. Summary

Perceptual landmarks at the workspace level will increase visual momentum by facilitating navigation between views. At a general level they are spatially dedicated (promoting consistency and automatic processing) and therefore facilitate the location of navigation-related information. They are powerful mnemonic devices, providing external landmarks that remind practitioners of destinations that can be navigated to and the associated functionality that can be accessed. They provide a visual “preview” of potential destinations (either new destinations or destinations previously visited) in the workspace. This preview orients practitioners, both visually and cognitively, thereby smoothing the ensuing transitions between views. Of course, perceptual landmarks in the interface must be both perceptually salient and distinct (relative to other landmarks) if they are to be effective.

The role of landmark knowledge (Lynch, 1960; Siegel and White, 1975) in navigation has been studied extensively. For example, May and Ross (2006) investigated landmark-based navigation aids in a driving field study; they found significantly better performance when high-quality landmark information was used. Empirical studies have indicated that the extensive research on navigation in real worlds generalizes to computer settings (Darken and Sibert, 1996; Ruddle et al., 1997). Vinson (1999) reviews the literature on the use of landmarks in these settings and provides guidelines for their design. Aretz (1991) found that a navigational display based on visual momentum principles provided a global set of landmarks that facilitated pilots’ abilities to orient to and locate terrain-based targets. The utility of landmarks has also been demonstrated in more traditional HCI applications. For example, Watts-Perotti and Woods (Watts, 1994; Watts-Perotti and Woods, 1999) found landmarks (e.g., table headings and
other forms of formatting) to be an essential design feature in supporting effective spreadsheet navigation (in fact, they observed that practitioners manually added landmarks when they were missing).

5. Overlap: rooms, center-surround, side effects, integrated representations

Woods (1984) originally described the overlap technique as a very general strategy to increase visual momentum by providing supplemental viewing contexts to frame the presentation of information. These contexts are presented in parallel and are used to testify with regard to physical contiguity, functional similarity, or other meaningful relations. For example, in cartography an area of interest (e.g., the state of Ohio) is presented in a detailed map situated within a larger, abbreviated context (i.e., partial maps of the surrounding states). Subsequent publications (Watts-Perotti and Woods, 1999; Woods and Watts, 1997) describe variations of overlap referred to as “rooms,” “center surround,” “side effects,” and “integrated representations.” Brief examples of each technique will be provided, followed by a summary.

5.1. Example: rooms

The rooms technique is designed to increase visual momentum by addressing a fundamental challenge often faced by practitioners: configuring and accessing sets of information or other resources that are required for the completion of multiple, asynchronous activities. Clumsy information systems, in combination with limited display real estate, often make this a daunting task (e.g., the command center described earlier). The everyday challenge of windows management (e.g., organizing, layering, resizing, opening, closing, moving) that often results in “window thrashing” (Henderson and Card, 1986) is a familiar example. Practitioners often adapt existing interface resources in an attempt to meet these needs (e.g., Cook and Woods, 1996; Watts-Perotti and Woods, 1999). Visual momentum can be increased by providing interface resources that allow alternative computing contexts or environments to be configured and accessed.

Some of the key features of the rooms technique will be illustrated using Apple’s Spaces® operating system feature. Fig. 5a represents a desktop environment containing a set of windows, applications, files, etc. (i.e., the current configuration, Work Context 1); assume that the practitioner has previously defined two additional work contexts (with different sets of windows, applications and files). When activated, Spaces provides a parallel overview of the workspace (i.e., all available work contexts) that is organized in a matrix (in this instance a 2 by 2 matrix with each work context located in a cell, see Fig. 5b). One way the practitioner can navigate between contexts (i.e., re-establish a previously defined work environment) is by pointing at and clicking on a cell in the matrix. For example, a click on the upper right cell in Fig. 5b causes Work Context 2 to become current (Fig. 5c).

Alternatively, practitioners can also navigate by activating a long shot display. This display provides an abstract, simplified version of the matrix with highlighting applied to the cell that corresponds to the current work context (e.g., Work Context 2 is active in Fig. 5c; it is located in the upper right corner of the matrix, as illustrated in Fig. 5b). Navigation between alternative work contexts is accomplished via arrow keys. For example, pressing the down (Fig. 5d) and left arrow key (Fig. 5e) in succession produces navigation from Work Context 2 (Fig. 5c) to Work Context 4 (Fig. 5d) to Work Context 3 (Fig. 5e). Overlapping elements (e.g., windows) within a work context can also be spatially separated instantaneously (via Apple Exposé®), thereby making them available for individual inspection (see Fig. 5f). This feature also works at the workspace level (e.g., all three of the work contexts illustrated in Fig. 5b would have spatially separated windows).

Woods and Watts (1997) provide a summary of design goals for the rooms technique. The first goal is to provide the potential to define alternative work contexts that can be tailored to meet task-specific needs (e.g., up to 16 work contexts in Spaces). A second goal is to provide a coordinated workspace (i.e., explicit representation of the relationships between work contexts and the ability to navigate between them, as exemplified by the Spaces matrix). A third goal is to allow work contexts to be seen in parallel (e.g., the matrix and the long shot displays of Spaces). A fourth goal is to allow practitioners to “… compose, save, and manipulate sets of views [work contexts] as a coherent unit…” (Woods and Watts, 1997, p. 644). Spaces supports this goal in a variety of ways. Composing a work context can be achieved via simple (i.e., starting an application in one of the desktops) or more innovative means (i.e., dragging an application or window from one work context in the matrix to another). Entire work contexts can be manipulated (i.e., point, click, and drag an entire context to a different cell in the matrix) as a way to organize the matrix along functional lines (e.g., work contexts with similar purposes can be placed close in space).

Early research projects addressing these needs include Smalltalk Projects (Goldberg, 1984), Rooms (Henderson and Card, 1986, an early and comprehensive effort after which the technique was named), 3D Rooms (Robertson et al., 1993), and Elastic Windows (Kandogan and Shneiderman, 1998). More recent examples include the Task Gallery (Robertson et al., 2000), the TimeSpace system (Krishnan and Jones, 2005), and Piles Across Space (Wang et al., 2009, for hand-held devices). Empirical evidence has been obtained in support of overlap techniques. Information visualization and windows management systems (e.g., Rooms) have been evaluated both informally (e.g., Krishnan and Jones, 2005; Robertson et al., 2000) and formally (e.g., Kandogan and Shneiderman, 1998; Wang et al., 2009) with generally positive results.
5.2. *Example: center-surround (focus + context)*

The center-surround design technique is used to increase visual momentum through parallel presentation: information in the practitioner’s current focus of attention is presented in high resolution (i.e., “center”); related information is presented in lower resolution (i.e., “surround”). The HCI/Computer Science community has used the label “focus + context” to describe this technique. Note that the center surround technique is conceptually similar to the long shot technique; one often-cited difference is that the center and surround are spatially integrated (i.e., they form a single integrated display) while the overview and detail displays are spatially separated.

The integration of center and surround is often achieved through a variety of systematic visual transformation and magnification functions (e.g., Leung and Apperley, 1994). Several representative examples will be illustrated via comparisons to a uniformly spaced rectangular grid of dots (Fig. 6a).

The “bifocal display” (Spence and Apperley, 1982) provides a center region where information is magnified and surround regions with uniform compression (Fig. 6b). The “table lens system” (Rao and Card, 1994) provides uniform compression in both the horizontal and vertical dimensions (and support for multiple centers and surrounds, see Fig. 6c). The “perspective wall” (Mackinlay et al., 1991) applies continuous, rather than discrete, compression (based on perspective geometry) for two surround regions (Fig. 6d). The “fisheye view” (e.g., Furnas, 1986) applies circular compression with continuous distortion towards the periphery. Fig. 6e illustrates a generic fisheye transformation using polar coordinates (e.g., Sarkar and Brown, 1994). The “document lens” (Robertson and Mackinlay, 1993) applies a modified fisheye transformation using Cartesian coordinates (Fig. 6f). Lampings and Rao (1996) also describe a modified fisheye transformation using compression based on hyperbolic geometry (not illustrated).

5.2.1. *Side effect views*

Note that while the information presented in a center surround display can be spatial in nature (and often is), other types of relationships can also be portrayed. Woods
Fig. 6. Abstract representations of various forms of information compression used in the implementation of the center surround technique. See text for details.
(Watts-Perotti and Woods, 1999; Woods and Hollnagel, 1987; Woods and Watts, 1997) propose a variation of the center surround technique that portrays functional relationships. Because of the many-to-many functional mappings that exist in certain work domains, the actions of practitioners will propagate through the system and produce side effects. These side effects can be either intended or unintended (and either for good or bad). The “side effect views” technique increases visual momentum by ensuring that practitioners are aware of these consequences (Woods and Watts, 1997, p. 646): “If an operator is interested in one function [i.e., the center] ... the system also provides a summary of the status of other functions which could be influenced by changes in the primary area of interest in a side effects window [i.e., the surround]”.

The need for side effect views is readily apparent for work domains, which have semantics that are driven by the laws of nature and therefore tightly coupled (e.g., process control). Woods and Hollnagel (1987) provide an extensive discussion of side effects in this category of domains and describe a control room display system containing a side effects view. This need is not restricted to process control; it is relevant for any work domain in which actions have propagating effects. For example, Woods (1994) proposed the use of side effect views to illustrate computational interdependencies within spreadsheets and between sets of spreadsheets.

Cockburn et al. (2008) provide an extensive review of the center surround (focus + context) literature. They conclude “It therefore seems clear that supporting both focused and contextual views can improve interaction over constrained single-view software” (p. 2:26).

5.3 Example: integrated representations

Woods (1984) originally described functional overlap as a way to increase visual momentum “By identifying [functional] relationships between data and user tasks and by paralleling those relationships in the structure of the display system...” (Woods, 1984, p. 237). Functional overlap is implied, if not explicit, in all of the overlap techniques described thus far. Subsequently, Woods and Watts (1997) identified “integrated representations” as a research trend related to visual momentum, but outside the scope of their review. Based on our experience in developing these representations (e.g., Bennett and Flach, 1992, 2011; Bennett et al., 1997), we believe that they provide a particularly useful form of functional overlap and that they qualify as a technique to increase visual momentum.

Cognitive systems engineering has developed analytical tools that have provided a more precise definition of functional overlap. The abstraction hierarchy (Rasmussen, 1983, 1986) provides five levels (categories or perspectives) that can be used to describe a work domain. Each level provides an independent category of information, yet there are inherent and systematic relations between these categories. Using this tool to model a work domain (e.g., Vicente, 1999) produces a description of the informational content (and inherent relations) that needs to be made apparent for a practitioner to think productively about complex problems.

Fig. 7 illustrates a representation that integrates information across levels of the abstraction hierarchy. The process represented by the display involves a liquid coolant entering (mass in), being stored in (reservoir level), and leaving (mass out) the system. The level of “functional purpose” contains descriptions of the design goals for the system. These are to maintain both a level of mass in the reservoir ($G_1$) and a flow rate for mass out ($G_2$). The level of “abstract functions” contains descriptions of the physical laws that govern system behavior. For example, the change in reservoir level (AR) should be determined by the difference between the rate of mass flowing into ($I_1 + I_2$) and out of (O) the reservoir. The “generalized functions” level refers to a source, a store, and a sink for fluid in the system. The level of “physical functions” contains descriptions of physical limits (e.g., the moment-to-moment values of each variable) and control (e.g., changing a valve setting). Finally, the level of “physical form” contains descriptions of the physical configuration of the system.

All of the levels of abstraction (except for physical configuration, not critical for control) are represented in this display (see labels in Fig. 7 and Bennett and Flach, 2011, for a more complete description of the design rationale and benefits).

The implications for increasing visual momentum are clear: if these relationships can be integrated into a single display, then the requirement to navigate between windows and remember information across these contexts is eliminated. Thus, the visual integration of information across the levels of abstraction increases visual momentum because it provides a “continuous graphical explanation” of events in the domain that is available in parallel (i.e., powerful perceptual skills are leveraged).

These integrated representations have been referred to as both “configural” (e.g., Bennett and Flach, 1992, 2011; Bennett et al., 1997) and “ecological” (Rasmussen and Vicente, 1990; Vicente and Rasmussen, 1992) displays in the literature and they have been evaluated extensively. Vicente’s DURESS system (Vicente, 1991) in particular has been systematically evaluated (e.g., Pawlak and Vicente, 1996). In addition, several studies have empirically evaluated the impact of integrating information from various levels of the abstraction hierarchy into representations (Burns, 2000; Ham and Yoon, 2001a,b). In general, representations that integrated information across more levels (as opposed to missing levels or levels that were displayed separately) were found to be more effective.

In the introduction we outlined the need to support visual momentum at lower levels of structure in the interface (i.e., the view or form level). To reiterate, obtaining information from a display (or collections of displays on a single screen) involves visual transitions that take place over time (serially). Visual momentum can be increased at levels...
lower than the workspace by providing effective visual structures that guide successive fixations; thereby increasing the “smoothness” of these segues. For example, does the screen provide optical structure that links information that is meaningfully related, but spatially separated? Does the display provide levels of optical structure that highlight the inherent relationships between the different types of information that it portrays? Tufte (1990) discusses the general principles of “layering and separation” that can increase visual momentum at the view and form level. We (e.g., Bennett and Flach, 1992, 2011; Bennett et al., 1997) discuss more specific principles to integrate and balance the hierarchically nested visual information (i.e., global/local “emergent features” and low-level “graphical elements”) in analog geometrical displays.

5.4. Summary

In general, overlap techniques increase visual momentum by explicitly illustrating meaningful relations (e.g., tasks, contexts, physical and functional perspectives). These relations are presented in an overall context that organizes alternative views (e.g., the matrix provided in Spaces). The currently active view is situated, in parallel, within this viewing context. Together, this provides a form of external survey knowledge (“you are here” and potential destinations) that serves to increase visual momentum by eliminating the getting lost phenomenon. Additional detail on related information can be obtained by changing to an alternative view. These transitions are facilitated by the overall context (i.e., it provides perceptual and cognitive orientation that allows the practitioner to anticipate changes in view), as well as natural and effective control mechanisms. The end result is a “widening of the key hole” through which the practitioner can view the work domain.

6. Spatial representations: spatial structure

The final design technique (spatial representation, see Fig. 1) has the greatest potential to increase visual momentum. It is a general technique used to incorporate (or impose, in the case of non-spatial data) a degree of
spatial thinking (e.g., Gauvain, 1993), environmental
1994), spatial knowledge (e.g., Siegel and White, 1975),
successively in a variety of ways: cognitive maps (e.g., Kitchin,
with information technologies. Specifically, virtual spaces
develops this technique in the following passage:

Spatial organization translates the normative user internal
model into a perceptual map. The user sees, rather
than remembers, the organization of data in the system
and can move within the system just as he moves in an
actual spatial layout ... One spatial access technique is
to organize the data base as a topology and then provide
the viewer with a mechanism to move through the space
... inter-display movements can be conceptualized as
tours or paths through the space ... the user’s
perceptual and attentional skills ... can be supported ...
by constructing a conceptual or virtual space ...

This technique increases visual momentum by leveraging
powerful natural skills, using them to facilitate interactions
with information technologies. Specifically, virtual spaces
allow practitioners to navigate between the views of a
workspace much like they navigate through man-made
(e.g., a familiar building) or natural ecologies (e.g., from
home to work). The associated human capabilities are both
general and powerful and they have been studied exten-
sively in a variety of ways: cognitive maps (e.g., Kitchin,
1994), spatial knowledge (e.g., Siegel and White, 1975),
spatial thinking (e.g., Gauvain, 1993), environmental
cognition (e.g., Evans, 1980), and way finding (e.g.,
Hutchins, 1995). The spatial representation technique
has, to a large degree, been implicitly recognized as an
integral part of HCI (e.g., witness the ubiquitous use of the
term “navigation”). The technique is well represented in
several of the interfaces discussed throughout the paper
(e.g. BookHouse, RAPTOR, iPhone, iPad) and will there-
fore not be reviewed in any greater detail.

7. Summary

The topic of visual momentum covers an extremely
broad spectrum of issues in design, as the variety of
techniques and examples discussed in this review indicate. Even so, our coverage of visual momentum was
necessarily selective in nature. We consciously chose the
pedagogical approach of covering a smaller number of
particularly representative examples in detail, rather than
providing a comprehensive treatment. We did not cover
some techniques that could very well be considered as
examples of visual momentum (e.g., animated transitions
between views) and did not discuss some excellent exam-
pies (Guerlain, 2007; McKenna et al., 2008; Patterson
et al., 1999).

As this review indicates clearly, techniques to increase
visual momentum are often complementary in nature
and can be applied in combination (e.g., long shot and
fixed format data replacement in Fig. 1; long shot, and
zoom+pan in Fig. 2). This occurs to an extent that
sometimes made it difficult for us to assign an example
or an empirical finding to a particular category.

On the other hand, these techniques can also be
competing alternatives for increasing visual momentum.
For example, the long shot and the center surround
techniques are reasonably similar in purpose and the
choice between them may depend upon inherent tradeoffs.
The center surround technique involves an integrated
representation whereas the two components of the long
shot technique (e.g., overview and detail) are spatially
separated (imposing a potential cost for information
access). Conversely, the compression of information in
the center surround technique saves space, but these
distortions can also make information difficult to access
(unlike the long shot technique).

We had several goals in writing this review: to once again
reiterate the need to consider visual momentum as an
important consideration in the design of information
systems, to refine and expand Woods’ original ideas, and
to integrate related concepts and findings from interdisci-
plinary literatures. In the process it has become very clear
that there is no single, preferred technique to increase visual
momentum; choosing one particular technique over another
will be dependent upon factors that are very situation-
specific. Cockburn et al. (2008, p. 2:26) reiterate this point:

The current state of research fails to provide clear
guidelines, despite a recent surge in empirical analysis
of the techniques’ effectiveness. Results to date have
revealed that the efficiency of different techniques is
dependent on many factors, particularly the nature of the
users’ task.

There is an important implication that can be drawn
from these observations. Specifically, the successful appli-
cation of visual momentum techniques will depend upon a
deep understanding of the design context: “Developers
should study/analyze what views need to be seen in parallel ...
... Methods for cognitive task analysis that analytically
map meaningful domain contexts identify views that are
inter-related and need to be seen in parallel…” (Woods
and Watts, 1997, p. 643, emphasis original). Others and we
have made the same points with regard to interface design
in general (e.g., Bennett and Flach, 2011; Rasmussen et al.,
1994; Vicente, 1999). Although they are sometimes viewed
as unnecessary steps in the process of designing computer-
ized decision support, these work domain analyses are
prerequisites to effective implementation.

In closing, experience has shown that the “myth of
information technology” referred to in the introduction
has been “busted.” Larger screens and increased automa-
tion (e.g., “intelligent” interfaces, expert systems) often wind
up increasing the complexity of modern work environments,
rather than decreasing it. The keyhole effect and the getting
lost phenomenon have not disappeared with the introduc-
tion of these and similar technologies; if anything, their rate
of occurrence has increased. The techniques to increase
visual momentum that have been described in this review
are successful because they leverage powerful technologies to increase the cognitive coupling between practitioners and their work domains. They are available now (i.e., they do not depend upon future advances in technology to be implemented) and they offer the very real potential to increase overall system performance.

Acknowledgments

We would like to express our sincere thanks to David Woods, Chris Wickens, Debbie Boehm-Davis, and anonymous reviewers for valuable suggestions that contributed immensely to the overall quality of the final paper.

References


Vicente, K.J., 1991. Supporting Knowledge-Based Behavior Through Ecological Interface Design. Engineering Psychology Research Laboratory, Department of Mechanical Engineering, University of Illinois, Urbana-Champaign, IL.


