Illustrative Interfaces: Building Special-Purpose Interfaces with Art Techniques and Brain Science Findings

W. Bradford Paley

Digital Image Design Incorporated, 170 Claremont Avenue, New York, NY 10027; Columbia University Department of Computer Science,
1214 Amsterdam Avenue, New York, NY 10027
`brad@didi.com`
`http://didi.com/brad`

Abstract. Restricting the scope of a user interface allows designers to apply domain-specific and task-specific knowledge, making the interface itself an illustration of the process it is meant to control. Such interfaces may be more easily learned and understood, and perhaps more efficient. This position paper demonstrates how we might apply to interface design findings from the study of human information processing and behavior, from psychophysics to behavioral psychology. It also explores the rationale behind using techniques derived from the fine and graphic arts. It suggests a simplified model of the mind as a "designer's checklist" that might help interface designers take advantage of the power inherent in people's ability to understand illustrations.

1   Introduction

Think of a point in history where technology can provide a single source for almost any kind of information—a channel for culture, education, entertainment, morality, law, and art. One might need to be a trained specialist to use it, but having it in your community (or even better) your home could begin an information explosion that would prove to be a major discontinuity in world culture.

I could be writing about the Internet-connected computer in a modern home, but I'm actually referring to the technology of printing in the years after its widespread deployment, starting with the Gutenberg's famous 42-line Bible, c. 1455. [1] The grand convergence of information in one physical interface wasn't the blessing of the new technology, but the curse of its immaturity.

As printed information technology matured we went from that single volume that needed a trained specialist to interpret—someone who could read—to the incredibly wide variety of form-factors it assumes today. Novels have different sizes and feels, different papers and different fonts than newspapers; labels on spice racks differ greatly from billboards. All for the simple reason that they are used for different purposes and today's well-developed print technology allows us to carefully match the "device" to the need.

Computers are just beginning to see this sort of physical articulation—calculators and digital watches prefigured personal digital assistants and multi-purpose cell phones by decades. And we are seeing many more unique devices proliferate, if not succeed. The whole field of Ubiquitous Computing addresses the expanding set of physical niches that we can expect to fit hardware information technology.

I suggest a parallel development in the scoping of computer interfaces, whether physical or merely graphical, which identifies specific task or domain niches and allow us to match the interface more closely to the idiosyncratic needs of each niche. As computing technology matures, more and more interfaces will become as specific as calculators and watches. This position paper explores techniques I believe will help make those interfaces more successful, techniques drawn from the arts and supported by an understanding of how the human brain's "protocol" allows artists to direct our attention and give us information.

1.1   Benefits of Restricting the Scope of an Interface

Interfaces that deal with very restricted scopes can be made much more specific than interfaces that must deal with broad ranges of information or many varieties of tasks. They can apply more domain-specific information and take advantage of lexicons and metaphors shared by a small target audience. When an interface looks and feels more like the task that's being accomplished, people can focus more on the task than the tool. The interface can become more "invisible" even if the graphical presentation is more elaborate.

In fact, a more elaborate or less common graphical representation can actually help someone use it more effectively. For example, a spreadsheet program can operate as both a calculator and clock, but seems more awkward to use for either purpose than the "real thing." I suggest that this is true partially because the consistent "numbers in boxes" look of a spreadsheet requires people to do more intellectual work. Realizing that the general purpose keys on a keyboard can serve to enter either passive symbols into a text or active numbers (textual representations of a potentially changing variable) may be an effort hidden from consciousness, but it still requires maintaining and switching to a distinct mental context when those keys used for a spreadsheet instead of a calculator. This mental context, or frame, might be more easily invoked if it can be associated with pre-linguistic, even pre-attentive memories or learning.

The hands on an analog watch have a very specific visual presence. Someone who wants to think about time might be more easily able to mobilize mental forces dealing with time if he's put into that mental frame by visual recognition of the hands on a clock. Physical stimuli can be strongly associated with memories, as Proust's madeline [2] famously demonstrates by sending him into an eight-volume reverie. And they may be just as strongly associated with active mental processes if the physical stimulus brings a whole mental frame including those active processes, to conscious availability (such as figuring out whether you're late or not, or subtracting two times). As interface designers we would like to see concrete evidence from psychological experiments demonstrate that this actually happens, but the lack of such evidence should not curtail design experimentation and innovation. In fact, experimentation in design can help us form new hypotheses for the sciences to test.

1.2   A Simplified Model of the Brain's "Input Protocol"

This position paper will suggest a few plausible connections that have been inspired by writings of people in the brain sciences. But they are presented as a simplification of the real processes, a kind of "designer's checklist" created to explicitly to support the human communication aspects of the process of design. Euclidean Geometry is a model of space proven to be incorrect, but carpenters still get useful work done with it. So our simplified model of the mind might help designers apply and invent new methods because it's easy to apply while we're engaged in the act of architecting information. And it can help individual designers to use some channels into the mind that they may be less familiar with.

A model, even a simplified one, is less useful without some means to apply it to the task at hand. That's where we look to another group of people who have been experimenting with directing attention and representing information for centuries. Artists and graphic artists have been researching human perception in a more ad-hoc way, but fortunately for the designer it's a very applied way. Their techniques can sometimes be disentangling into the myriad individual neurological and psychological processes that combine to, for instance, direct the eye one way or another. This helps us build our model of the mind. But they can sometimes be lifted whole from a painting and applied as an integrated technique. This helps us build interfaces.

The techniques of artists have helped shape our model by providing questions: "why does this mark make me look there?" They also show where simplification is acceptable, letting us group dozens of distinct brain processes into six or nine general categories. Future work includes making a list of artist-derived interface techniques, how they are described by our model, and what sorts of tasks and information they might support. But a listing of those techniques is well beyond the scope of this position paper.

So this paper will outline our still-developing model, a "designer's checklist" that helps describe how people process information, with just enough simplification to make it useful for design. And a technique or two that may have been originally drawn from artists' palettes will be used to demonstrate different parts of the model.

Many of these techniques may have application in building interfaces for general-purpose data exploration or document creation. But the best of them get their strength from tying abstract bits of data to meaning in someone's head by illustrating that meaning—by crafting a user interface widget that has as much of the character of the idea as possible, so it's difficult to mistake for something else. Future work also includes sorting through this growing set of techniques to see which can be applied to all interfaces, and which are best tied to a restricted task domain. But for now the arena of special purpose interfaces seems to be richer in suggesting candidate techniques.

2   A "Knowledge Acquisition Pipeline"

People turn marks on paper and patterns of light on computer screens into living, active ideas in their minds. How does this happen? Contemporary brain sciences have identified dozens or hundreds of distinct neurological processes that participate in this remarkable transition. As interface designers, we can group and simplify them into a

few categories that help us do design. In creating this simplified model we should keep the purpose of the model in mind. We're trying to support the design process—to help us figure out how ideas come from information displays and how those displays can suggest actions that will help people modify the information.

The following breakdown has proven useful for me to organize my own exploration of what kinds of displays might evoke what ideas. It will undoubtedly improve with time, and I welcome thoughts from those reading this initial attempt. Let's use the visual modality for our examples, though this model might as easily be applied to others (e.g. auditory or tactile).

- The first step in the process of reading, for example, is a sensory input: the eye sees areas of light and dark; color variations of hue and saturation. Call this a *sensory* process.
- That raw sensory data feeds into processes that integrate or differentiate the sensory input—some starting immediately in the retina itself, such as the lateral inhibition of the retinal ganglion cells which helps us identify boundaries and see lines. Other processes identify things like line orientation and surface texture, the building blocks needed to segment a scene into pieces. Call these *perceptual* processes.
- Some of these lines and textures help us recognize objects, a *cognitive* process.
- Some objects, like the letters on this page, are associated with ideas beyond the object itself; a *symbolic* process.
- Symbols can be assembled into larger structures; like words, sentences, or topics; a *linguistic* process.
- And finally we can understand and commit to memory the propositions carried in language; a *semantic* process—the ideas we as information designers are trying to evoke.

This relatively cold way to think of a person ignores issues critical to careful interface design, such as behavioral, emotional, and social processes. Though they don't fit nicely into the simplified linear pipeline, they're valuable in our checklist, so at this point we just add them as free-floating reminders, to make our final interface designer's checklist:
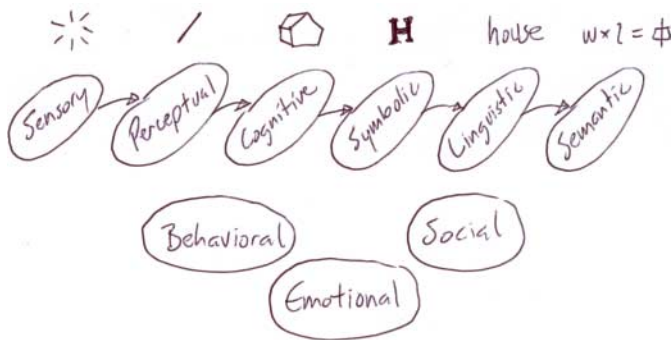

Fig. 1. A "Knowledge Acquisition Pipeline" and other reminders (Intentionally rendered as a sketch, to visually indicate work in progress)

Each set of processes can be thought of as a channel by which we can get information into someone's mind. [3] The most remarkable thing to be exposed by this breakdown is a serious asymmetry in how these channels are used. The vast majority of information in most current interfaces—especially those built with standardized tables, input fields, lists, and scrollbars—are decoded by people's symbolic and linguistic processes, very late in the process. It has to be decoded by the forebrain often with specific, concentrated attention. Even icons are symbols, and the colors used in many visualization displays must be processed as symbols, too—if you need a key, it's a symbolic mapping. Some specific applications; like computer games, computer-based art, and scientific visualization tools; have made better use of the earlier channels but they remain isolated oases of visual sophistication in a desert of text and rectangles.

The ironic thing about that is the forebrain is only three to five million years old, compared with some 125 million years of evolution behind the human visual system; and it's "single-threaded" since we can only attend to one thing at a time, compared to the hundreds or millions of things going on simultaneously in the earlier processes. [4] Worse, we may not even be using it for what it was evolved for: the convoluted logical manipulations needed in a lot of modern tasks developed well after many scholars think evolution stopped for humans. At this point in the maturity of computer interfaces we are squeezing huge amounts of data into a narrow channel: a johnny-come-lately, single-threaded, possibly misused fraction of our brain.

This asymmetry is exciting because it points to a gross inefficiency in how we currently work with computers, and suggests how to fix it. If we find ways to feed some of the data into earlier channels we can greatly improve the man/computer synergy. An intuitive recognition of this asymmetry may be what fueled the mass media's great hopes for virtual reality, and to a lesser extent for scientific visualization. I believe that we have barely delivered on the promise of these fields partially because it's difficult to thread meaning all the way through the pipeline, from sensory input to semantic value. The techniques of "Illustrative Interface Design" provide a first gesture towards exploring and codifying techniques that have been explicitly developed to make that end-to-end connection.

3   Example Techniques

In this section I'll describe applications of visual techniques that fall into the categories in our simplified Knowledge Acquisition Pipeline above. In this short paper I will concentrate on the first three channels: sensory, perceptual, and cognitive processes, since the final two channels are in more common use.

3.1   Sensory Processing Examples

Sensory processing is the earliest of the channels in our pipeline, and one of the oldest in evolutionary terms. Along with perceptual processesing, it may be the most effective in triggering what biologists call "orienting behavior," associated with the release of adrenaline, faster heart rate, and alertness. This state is well suited to

directing attention toward important events, where "important" must be determined by the business rules or prioritized task list of the person using the system.

Fig. 2. Detail from a design for an Equities Sales Trading system. A similar system has been working for three years on the New York Stock Exchange trading floor. [5]

In Figure 2, a detail from a design for an equities sales trader's workstation, let's look for the most striking visual features: features that will still be visible even if the image is blurred, shrunken, or displayed for a small fraction of a second. The default colors for the interface were picked to have a relatively bland and flat look so that we could make new information stand out. Here, newly arriving information is one of the most important things a sales trader needs to manage, so new things are distinguished from the rest of the interface by several visual sensory attributes: they are darker, higher in contrast, letters in them are reversed to light on a dark background, and colors are more saturated.

A quick scan of this interface will cause someone's attention to be drawn immediately to the fifth and eighth objects in the list, large areas accentuated by all of the above attributes. A finer-grained scan may draw attention to the first, ninth, and tenth objects, each of which has a smaller rectangle painted with these attributes.

The rest of the interface is intentionally less visually demanding to provide a quiet background for the important information. But information is still immediately available for the trader as soon as they look in the right place. For instance all of the other objects are closer to the background color, drawn with a medium level of contrast. The window borders, including the scrollbar, are drawn with the lowest contrast; appropriate for something that should draw the least visual attention.

In this interface the scrollbar is always there, so it should be available but barely visible—the trader always knows it's there to use so there's no need to call attention to it. It's ironic that in most contemporary windowing interfaces some of the strongest visual cues we can provide on a computer screen—contrast, color, and a 3D appearance—are being squandered; directing the user's attention *away* from the real information in the documents they're creating.
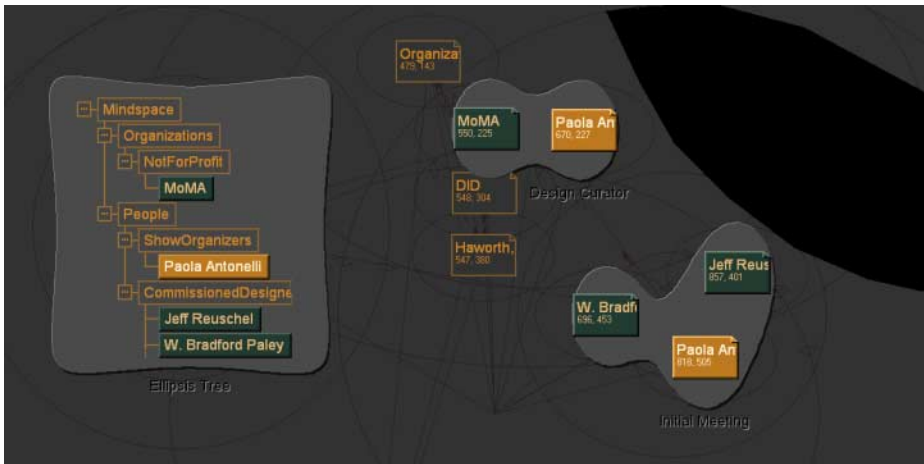
Fig. 3. Detail from a knowledge management concept design, commissioned for the 2001 WorkSpheres exhibition, by the Museum of Modern Art, New York [6] [7]

Figure 3 shows the same sort of attention-directing visual attributes—differences in lightness, saturation, and 3D appearance—at work in another design. This design supports two different views on the same database, or "Model" as described in the Model View Controller paradigm. [8] One view, a hierarchically structured view, is on the left, labeled "Ellipsis Tree." The other view covers the rest of the image and lets people make less structured, ad-hoc arrangements of the information objects.

The highlighting not only draws attention to the selected object (labeled "Paola Antonelli"), but it makes is clear that the object is represented in three places: one in each of two ad-hoc "blobs" and once in the hierarchy. When any object is selected, it and all of the other objects representing the same entity in the Model database acquire the most visually demanding attention-directing visual attributes: bright orange with 3D edges.

This method of highlighting can be distinguished from standard highlighting by its conscious use of low level sensory input. Typical highlighting schemes do not necessarily take into account the colors or attributes on the rest of the screen. A selection color is chosen and represents the locus of attention by a more symbolic mapping: selection color means user focus. Since the colors used in the rest of the screen are not rigorously controlled to provide a low-level sensory contrast with the highlight color, it may take an active change in the user focus to distinguish exactly what is highlighted and what is in the background. It may not be obvious from a still image, as it is here. A case sometimes seen in a standard text list illustrates the point: if half of the items are selected it's often not clear which half is selected unless you know the selection color, or watched (or did) the selection yourself. Your eye is not immediately told what's on and what's off.

It can also be distinguished from an unfortunate technique of using one visual attribute per information channel. Especially in early scientific visualization systems, designers became enamored of how much information could be written on the screen, with the goal of the highest data-density possible. This approach ignored the point of

many data representations: that the goal was *reading* the display, not writing it. Reading a display can be made much more difficult by treating visual channels that have deep links in our minds as separate ways in. For instance, having three different ways to highlight things represented as red, green, and blue channels in an RGB display does allow one to have orthogonal selection sets. But people don't experience yellow as R+G, so the complex overlapping of three selection sets becomes an even more complex task of decoding a symbolic proposition that allows eight variations of highlight and keeping it in mind over potentially disjoint spatial areas and thousands of objects. Visual attributes evolved in richly-connected webs of dependence and mutual support, and there is evidence that they are best used in concert [9], as the redundant-encoding of highlight state is accomplished with several visual attributes in Figure 2 and especially Figure 3.

In the MoMA system, a less-demanding color is also used as a clue to show what objects might be important to look at for the next step in the task: dark green, with lightness almost matching the lightness of the background, medium saturation, and medium-lightness labels. This two-level "Selected/Related" highlighting scheme is useful when the system has rules that can help people with the next step in their thought processes. Here, it is assumed that ad-hoc "blob" relationships are telling the system "whenever I select one of these objects, remind me that I have associated the other ones with it."

3.2   Perceptual Processing Examples

There are three kinds of lines in Figure 2: (a) an implied line between two areas painted different colors, (b) a simple line drawn on a background of a different value or color, and (c) the highlight and shadow lines used to create the 3D appearance of the "information objects." They are used to indicate and distinguish three different kinds of entities in the interface

The implied line (a) may be the least visually demanding (at least when drawn in the very close colors we use here). It is used for the interface elements with the least freedom—the background behind the objects and the edge between the background and the window borders. These edges serve to distinguish different areas, but can not be clicked or dragged; they are not affordances.

The simple line (b) is used to outline more active interface elements: the scrollbar, and the thumb and arrow areas that are parts of it. These are clickable or draggable affordances of the scrollbar, but can only be moved (if at all) within tight constraints. This is also true of the small vertical rectangle inside the objects: it acts as a clickable affordance that will bring up a "details window" for that object, so it is generally drawn with the same sort of line. (The exception is when the details window is dark to indicate new information, but in that case there's a more important message being given to the trader. It's okay to override our convention for distinguishing affordances to allow the more important cue to dominate visually.)

In Figure 3, perceptual cues are at work in one additional way: information objects are drawn with straight lines, and relationships are drawn with curved lines. Even before there's a distinction in the viewer's brain between objects and background, there's a distinction between the ways we encode data and relationships. I suggest that this may make it easier to understand the differences later on in the Knowledge Acquisition Pipeline: when earlier processes are supporting later processes, and the

signals all encode the same information, we have a very strong redundant encoding of that information.

The highlight, "low-light", and shadow lines (c) in both figures 2 and 3 create the illusion of 3D objects, leading us to the next step in our pipeline, the cognitive processes.

3.3   Cognitive Processing

The human visual system devotes a great deal of effort to finding the borders of things, splitting the world into objects and background. This may be because finding food, threats, and possible mates has a high survival value, and these goals are generally embodied in objects. It also helps us to identify things that we can grasp and manipulate. The cognitive process of distinguishing objects may be even more developed because we're constantly using our hands with nearby objects and the plasticity in the human brain allows it to devote more neurons to processes that are used more or require finer distinctions, even after nature applies the genetic blueprint.

To take advantage of this, we design information objects out of entities that are important to the trader rather than leaving them listed as abstract labels in a standard table format. Our rule of thumb: if something is an object in the trader's head it should look and act like one on the screen.

"Orders" are requests from a customer to buy or sell a financial instrument, the central thing in a sales trader's job, things that have powerful character and cohesion as entities. An order can be urgent, dangerous, unexpected, far out of the market, completely new, or partially updated with new information. Putting it into a standard table requires taking a living, articulated idea and disaggregate it into abstract attributes written as words or numbers in columns. This data normalization completely ignores the emotional and physical capabilities of the mind which help to distinguish, enhance, and identify objects and give them meaning to the trader. Each order must be mentally reassembled by scanning across the table, logically realizing that each row is an entity—despite the fact that a typical table makes no distinction between rows and columns, and despite the fact that columns may have a stronger visual cohesion than rows do since they're adjacent along the longer side of each rectangular cell. By drawing something as an object on the screen we may be making it easier to associate the information on the screen with the entity in the trader's mind.

There are two visual pathways in the human brain, called by some the "what" and "where" (or "how") pathways. [10] The *how* pathway is the earlier one in terms of evolution, and connects with the motor system to control our hands and motions when we manipulate objects. The *what* pathway connects to our conscious processing, letting us the understand things about an object. I suggest that when an object is rendered in 3D it might engage both pathways more firmly, making the 3D object easier to focus on and manipulate mentally as well as physically. In Figure 2 the scrollbar, title bar and resizing edges remain flat, while the order objects are 3D, since they are the usual target of manipulation.

In other systems, such as in Figure 3, I have left many information objects flat, as context for the few 3D objects that are needed for the current manipulation or thought process. Objects become 3D when they need to be manipulated, and the previously 3D objects flatten out to become the new context. The dramatic visual difference

makes it feel easier to focus on and manipulate the 3D objects, but actual efficiency increases have to be tested and proven.

In the sales trader's system the more an object "looks like" an order the easier it is for him to associate it with his mental order. In this design the order object's triangular end makes it read as an arrow, implying the "flow" of shares from one party to another. This clearly distinguishes sell (right-pointing) from buy (left-pointing) orders—a very important distinction to traders—while keeping them all looking like they're members of the same family. The visual "objectness" helps the trader's mind to be ready to manipulate this order as a whole and distinct entity, and the shape itself implies some of the actions that might be taken.

Likewise in Figure 3, ad-hoc relationships are fluid blobs that change when their member objects are moved, while the structured tree is more rectilinear. I suggest that this shape difference is a useful mnemonic that suggests the permanence and structure of the hierarchy and the transitory, less explicit associations in the blobs.

3.4 Symbolic and Linguistic Processing

Even the symbolic and linguistic information can be visually articulated in ways that make it more understandable and readable to the trader. In figure 2, for example, we draw the most important information in the largest fonts, and strike through numbers that are not operational but historical. At a lower level this font variation provides visual landmarks to help people locate specific pieces of data.

4 Illustrating Information

There are many ways to make an information object look like an entity in someone's mind, and the exploration of how to do this most effectively is one of the central issues in developing Illustrative Interfaces. Here, we've touched on a basic first principle—associate a visual object with a mental entity. More refined principles all developed along similar lines: the visual object should depict as much of the character of the mental entity as possible. This is where we start seeing the luxuries a designer has in a restricted application scope: more restricted scopes often have more of the idiosyncrasies and details that can be represented in a more evocative way.

To accomplish this while designing an interface for an individual or group, we can first make a list of their language, structures, notational conventions, and shared metaphors. Then we can try to illustrate them: make some kind of drawing (or sound, or tactile event) that captures some of the character of the mental entity by using their own language or notation.

Specific techniques we use to capture character vary widely.

- We can simply copy previously used artifacts to the screen. This is very useful because people recognize them. It may be one of the reasons the Quicken personal finance system became so popular: there are concrete recognizable checks on the screen, not just numbers in columns.
- We can simplify existing objects and redraw them.
- We can draw completely new objects based on the entity distinctions in the minds of the group. And we can shape and decorate them to display

> information in a way that resonates with how people think of them—a kind of applied "visual data poetics."

This last technique is the one demonstrated in the sales trader's system example in Figure 2. I believe it holds great promise for showing the widest range of information, since drawn information objects can be shaped to take advantage of the fluid representational abilities of the computer. Once people have a concrete representation of their current thoughts in front of them their minds are freer to build new levels of ideas on that foundation. These new ideas may suggest another round of information objects, acting as bricks in an iterative climb towards more complex yet more understandable structures of ideas.

5   Conclusion

I suggest that the use of an interface designed to reflect people's thoughts about their work, leveraging the illustration-interpreting capabilities of the mind might have several very positive effects: It might speed learning of the system by people in the group. It might help a group teach and integrate new colleagues. It might help a group share a stronger, more concrete mental model of the activity they're all engaged in. And perhaps the most profound effect is that if that group's shared mental model is flawed, its new concreteness might make it easy to identify the flaws and focus on how to repair them or create a wholly new mental model.

This new mental model might be useful in itself, as a new scientific understanding. Or it could be the basis of the creation of a new work process, with fewer errors or costs, or greater efficiencies. Of course then we have to design a new interface, but iteration is the heart of any good design process—this approach to interface design just might make it easier to include the user's own work or though process in the iterations.

References

1. Blumenthal, Joseph: Art of the Printed Book 1455-1955, Godine, Boston, MA (1973) 2
2. Proust, Marcel: Remembrance of Things Past, Volume I, Vintage Books (1982)
3. Kosslyn, Stephen M., Elements of Graph Design, W. H. Freemen & Co. (1994)
4. Minsky, Marvin, The Society of Mind, Simon and Schuster (1986)
5. Abrams, Jan, Stock Options, ID Magazine, June 2000, F&W Publications, New York, volume 47, number 4, (2000) 48–51
6. Paola Antonelli, Ed., Workspheres, Design and Contemporary Work Styles, The Museum of Modern Art, New York (2001) 215
7. Paley, W. Bradford, Web site with MindSpace production notes and working Java applet, http://didi.com/brad/mindspaceAtMoMA (2001)
8. Glenn E. Krasner and Stephen T. Pope. A Cookbook for Using the Model-View-Controller User Interface Paradigm in Smalltalk-80. Journal of Object-Oriented Programming, 1(3):26–49, August/September 1988.
9. Zakia, Richard D., Perception & Imaging, Focal Press, Woburn, MA, (2002) 61
10. Ramachandran, V.S., Phantoms in the Brain: Probing the Mysteries of the Human Mind, Quill, New York (1998) 77–82