Query by Attention: Visually Searchable Information Maps

Mark A. Foltz and Randall Davis
MIT Artificial Intelligence Lab
Cambridge, MA 02139 USA
{mfoltz,davis}@ai.mit.edu

Abstract

This paper explores how the design of information spaces might be grounded in knowledge of human visual processing, notably what kinds of visual selection are most efficient. Information maps spatially array graphical symbols representing items of information and their attributes. Ideally, their users should be able to do query by attention: *answer questions about the information quickly by controlling visual attention (i.e., through spatial selection and visual search), instead of manipulating an interface. I propose a preliminary method for designing visually searchable maps based on experimental results about what kinds of visual search are easy. The hope is that the resulting maps will better employ the perceptual capabilities of their viewers when they search. An example information map of recent movies illustrates the approach.*

Keywords: *information maps, presentation design, visual attention, visual perception, visual search*

1. Information Maps and Visual Search

Reading a map like that in Figure 1[1] to navigate the Boston subway requires at least two episodes of visual search: find the originating station, then find the destination. Could we improve the design of the map by using our knowledge of what kinds of visual search are easiest?

A small visual stimulus that appears to be at a different depth relative to its surroundings "pops out" – it can be found almost instantaneously [27, p. 39]. We could add such a stimulus to a map posted inside a station indicate that station (Figure 1, right). Now a map reader (with an appropriate legend) can locate the station by efficient visual search, instead of scanning the text labels.[2]

[1]Color versions of the figures are available at `http://www.infoarch.ai.mit.edu/publications/`.

[2] Using Wolfe's classification of search task efficiency [27]. A classic example of efficient search is locating a single `X` among `O` distractors. Search time increases with display size with a very shallow slope ($<$ 5 ms/item).

Designers of information spaces would like to make the same kinds of improvements to more complicated displays. This paper explores the relationship between information space design and visual search, in the hope that spaces can be made that facilitate rapid visual queries. The preliminary proposal is that they should be designed to make maximal use of spatial selection through attentional control, and visual properties that pop out, i.e. require almost no effort to find even in a display with many items.

In this paper, an information map is a display that spatially arrays graphical symbols representing information items and their attributes in two dimensions. By *information item* we mean an object with attributes meaningful to users, e.g. a movie, document, or historical event. The task is to obtain the subset of items whose attributes match the query criteria. For example, in the Beethoven symphony map in Figure 2, we can find which pieces were composed before 1808 by attending to the upper third of the map. Or, we can find the symphonies in minor keys by attending to the textured symbols. Once attention has been directed to the subset of the items, further inferences can be drawn about them (e.g., no symphony in C was composed after 1808). (This map is similar in layout to GlassEye, a visualization of Phillip Glass' works [14], but was designed independently.)

1.1. Why information maps?

Information designs that use the position of graphical primitives have long been used to concisely convey data [3]. An advantage of these designs is that the user can *query by attention* – answer questions by controlling visual attention (and receiving immediate feedback), rather than manipulating a database query interface and waiting for a server to return results. The user can more rapidly explore the information space because he needs to alter only his visual attention (and not the interface) to adjust the parameters of the query.

Information maps, which give meaning to a item's location, also engage their viewers' capabilities for spatial im-

Appeared in *Proceedings of Fifth International Conference on Information Visualization (InfoVis 2001)*, pp.85-93. 2001

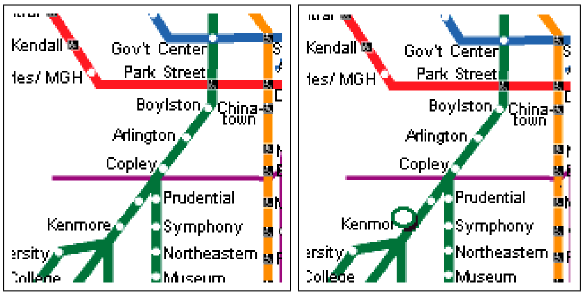

Figure 1. Left, part of the Boston subway map. Right, an improvement using a depth cue to highlight the nearest station.

agery and spatial memory. If the user remembers an item's location, he can return to the map and find it there, instead of repeating a textual query or scanning a list. With the right design, information maps are a less cumbersome tool for querying and drawing inferences from the ever-growing amount of information we are faced with daily.

Information maps have limits: there are a limited number of independent visual properties that support rapid visual search. And, using too many properties has its drawbacks. A map that uses its symbols' position, color, shape, depth, *and* orientation to encode information risks overloading its viewer's ability to attend only to the properties relevant to his information need, by the presence of multiple distracting properties. Another risk is that the mapping of properties to meanings must be retained in the viewer's limited working memory.

To get around the problem of a limited number of visual attributes, maps can be interactive (displayed on a computer). The user can obtain multiple views of the same information by selecting relevant attributes to display. If there are too many items to display without crowding, he can prefilter the information in the map [1] or browse it with a hierarchy. These methods help to bridge the gap between query by attention and query by manipulation.

2. Designing for Query by Attention

Consider the common scenario in which a user wants to explore a small database by posing a group of related queries (a task that would ordinarily require several transactions with a textual interface). Assume that the database is small enough that a map can display all the items without overcrowding.

The information is treated as a bag of items; adding or removing an item would not substantially change the overall meaning of the map. Each item has attributes, which are variables that can be binary, enumerated (chosen from an unordered set), integer, real, or textual.

A *query* is a set of restrictions on the values of an item's attributes that must all be satisfied for an item to match. An example of a query in Figure 2 is `Key = F AND Year < 1810`. The user explores the information space bottom-up – adding and removing restrictions to a working query and seeing which items match.

These maps are a simple kind of information space that set the context for this initial sketch of the constraints and affordances that relate visual search to information space design.

2.1. A Design Method

We propose a four-step method for designing a visually searchable information map:

1. List the queries viewers are likely to make and the information attributes that are available to answer those queries.

2. Propose a mapping of information attributes to visual properties, using the design flowchart (discussed below) to suggest which mappings can be searched efficiently.

3. Produce an initial map and test it to see if users can quickly and easily make the kinds of queries listed in Step 1.

4. If some of the queries are slow, adjust the mapping of information attributes to visual properties and repeat Step 3.

Of course the real difficulty in this method lies in Steps 2 and 4 – finding the initial mapping and adjusting it so that the map is usable in practice. The design flowchart discussed below constrains the possible mappings in terms of attribute types, but does not specify a unique mapping. Instead, this approach is iterative: using what we know about the cognitive affordances for search, try an initial design with these guidelines, then iteratively improve it with feedback from users. A good design solution will match the user's patterns of semantic attention across the information with patterns of visual attention across the map.

Another difficulty is that the number of relevant attributes to display often exceeds the number of available visual dimensions. In this case, the designer must consider which attributes best support the user's query goals. (Dimension reduction can also help, as discussed below.) Finding the best attributes often requires user input, iterative prototyping, and careful task analysis [12], which are beyond the scope of this paper. Here we assume we are given a data model and consider how to display it in a visually searchable form.

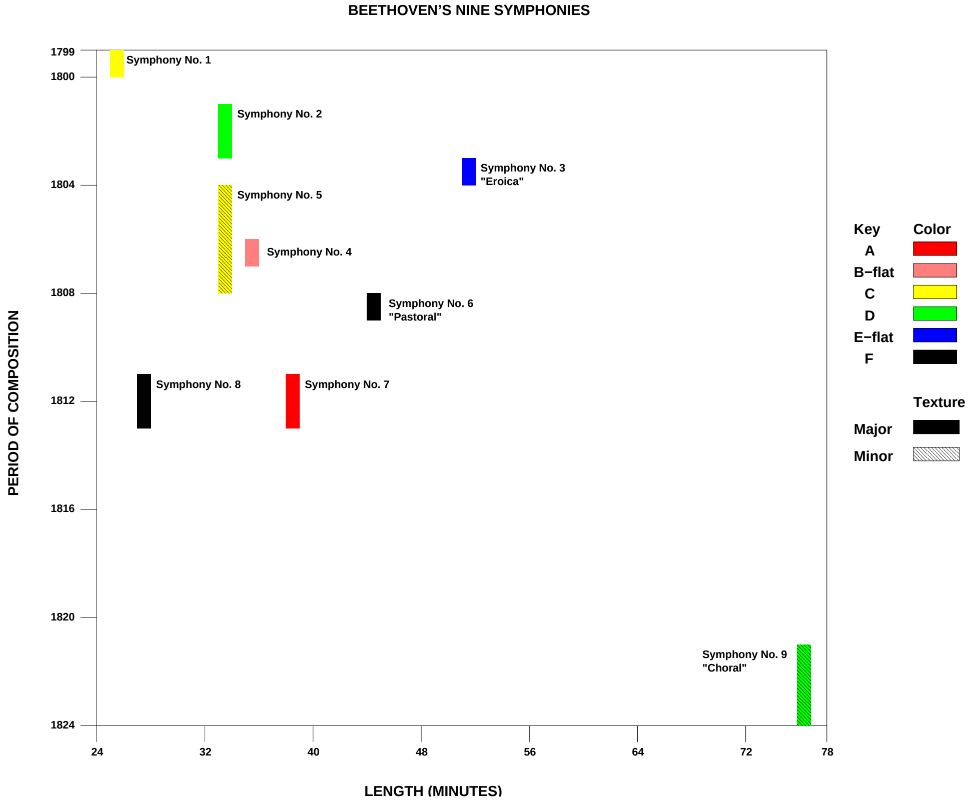

Figure 2. An information map of Beethoven's nine symphonies. Color and texture are used to indicate the key of the composition.

3. Visual Properties

This section considers in more detail the visual properties available to represent a symbol in the map. We begin with single visual properties that can define a symbol's appearance, then consider combinations of those properties that can support efficient conjunctive search. This summary is based largely on Wolfe's review of visual search experiments and results [27], and Feature Integration Theory [25].

3.1. Single Properties

Research has shown that many visual properties can support efficient visual search. Below we consider visual properties that have been used to convey information in print information design [3, 26], and are also discussed in Wolfe's review.

Position. A user can control visual attention to search within a spatially delimited region, making it appropriate to map to the X- or Y-axis attributes that are queried by range (e.g. *Find all films in the 1980s*). The user can quickly shift attention to different scales or regions, easily adjusting the query.

An axis can also be subdivided to allow one spatial dimension to encode more than one attribute. (This approach has been independently applied to graph multivariate functions [17].) A primary, enumerated attribute can partition an axis into subaxes, and a secondary, continuous attribute can be mapped onto each subaxis. For example, we could add Beethoven's contemporaries to Figure 2 by partitioning the X axis by composer, then plotting symphony length on each subaxis. Although subdivision allows more attributes to be represented at once, searches that require the secondary attribute but not the first require the serial examination of several disjoint regions of the map.

Color. Color is widely used to distinguish symbols in maps. The colors at the extrema of opponent processing [19, p. 113] (red/green, blue/yellow, and black/white) are considered good candidates for efficient search. D'Zmura proposes that a color will pop out if it is linearly separable in CIE color space from distractors [9], while other work has shown that search for a color target among as many as nine distractor colors is efficient if the colors are spread far apart in color space [22].

This suggests choosing colors that are symmetrical in a ring of high saturation in color space, maximizing distance and separability (Figure 3). If white is reserved for the background, then at least five colors are left for map symbols. Color is best for enumerated attributes that can take two or more values. [3]

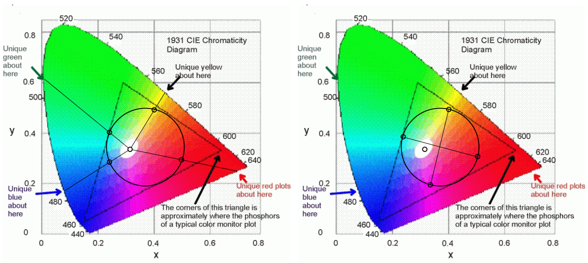

Figure 3. Left, a set of symbol colors (open circles) chosen to approximate the extrema of color opponent processing (indicated by arrows). Right, colors chosen to maximize separability. The triangle encloses the colors displayable on a CRT monitor.

Shape. Because there are many parameters that determine shape, coming up with a fixed set of rules for finding shapes that can be searched for efficiently is difficult. Some determining features include line termination (presence/absence), closure, holes, and possibly intersection [27, pp. 31-4]. Most experiments described by Wolfe report efficient search with homogeneous distractors, so a conservative approach would map shapes like `X` and `O` to a binary attribute. Orientation also supports efficient search, for example using bars oriented at 0 or 90 degrees.

Motion. Motion exhibits search asymmetry [21] – a moving target among stationary distractors is easy to find, while a stationary target among moving distractors is not. Thus, a map that uses motion to code a binary attribute is biased against searches for the stationary value. However, the user may choose to render a subset of the symbols as salient in an interactive map, and search among them. This prospect is appealing, because search for motion in conjunction with other properties can be efficient (as discussed below).

Depth. Pictoral depth cues such as shading, occlusion, or shadows can support efficient search [27, p. 39]. This cue is best used for binary features, as search in more than two depth planes is less likely to be efficient.

Other cues. Wolfe lists other basic features that can be searched for efficiently: vernier offset, curvature, gloss, size, etc. While these are also worth investigating as useful for information maps, the cues discussed above – position, color, shape, motion, and depth – have been investigated thoroughly in the visual search literature. This initial proposal explores the possibilities they afford for information map design.

3.2. Combinations of Properties

The original version of Feature Integration Theory [25] held that visual properties are preattentively and simultaneously processed into feature maps across the visual field. Searching for a target defined by a single property is fast because only one feature map is consulted to locate it. To search a conjunction of features, however, information from multiple feature maps must be bound together at item locations, and each location serially checked to see if it possesses the target conjunction.

Later work has shown that for certain combinations of distinct visual properties, subjects can perform conjunctive searches efficiently. In particular, Egeth et al. as well as others have shown that subjects can efficiently search a colored subset of items for a differently shaped target [10]. Efficiency improves when the targets are far apart in color space.

Other work has shown that visual search is efficient for a moving stimulus among stationary distractors that also differs in depth, orientation, or shape [7, 8, 15, 16] (but as mentioned, motion search is asymmetric).

Stereoscopic depth can also be used with color or orientation for efficient conjunctive search [18]. Computer-generated stereo displays usually require special viewing equipment, however.

These results are important because they suggest a user can query by at least four attributes, first by spatial selection (using two real-valued attributes), and then by searching for a conjunction (e.g., color and shape).

3.3. Summary

The flowchart in Figure 4 summarizes the suggested design combinations that afford efficient search. A mapping is created by tracing a path from START in the graph, assigning visual properties to information attributes of the indicated types. The edge chosen to leave a visual property node assigns its information attribute (e.g., placing `SymphonyLength` on the **R** edge leaving **X** means that the map will have symphony length on its X-axis). Names or other labels may be ordered alphabetically and treated like integers in the flowchart, or included as text adjacent to map

[3] It is important to note that 8 percent of men and 1 percent of women are colorblind [19, p.104], so visualizations that use colored symbols should provide an alternate presentation of that information (e.g. with shape).

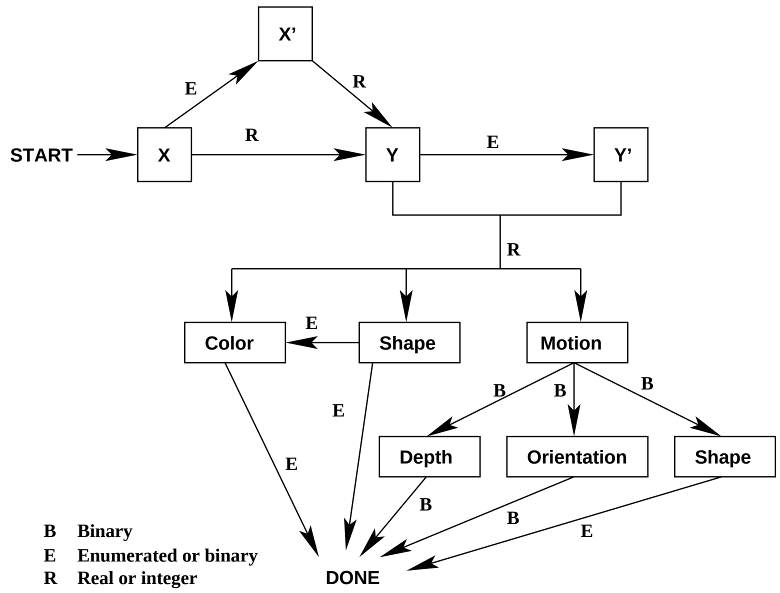

Figure 4. A design flowchart for visually searchable information maps. A map design traces a path from START, assigning visual properties to information attributes along outgoing edges. X' and Y' represent the subaxes created when the X and Y axes are partitioned by an enumerated attribute. The path may terminate before reaching DONE.

symbols. (For simplicity, the latter case is not included in the flowchart.)

4. Example Information Maps

To illustrate the method, we first consider the problem of designing an information map to visualize a database of about 25 movies. Suppose the database contains the attributes shown in Table 1.

Attribute	Type	Example
Title	textual	`Titanic`
Year	integer, 1900-2001	`1999`
Genre	enumerated	`Drama`
Oscar	binary	`Y`
Review	real, 0.0-10.0	`6.8`

Table 1. Movie database attributes and example values. The review scores are from Internet Movie Database votes.

The first step in the design procedure is to list the questions users might ask of the data. The questions in this list depend on the motivation of the prototypical user (e.g., a video rental browser versus a film researcher). Table 2 shows some sample questions that might be asked by someone browsing for good, recent movies.

In these sample questions, both `Year` and `Rating` are ordered attributes queried by range, so they are good candidates to map to axes. `Oscar` is binary and can be mapped to two shapes, while `Genre` is enumerated over five values and can be mapped to five colors.

Combining these observations, a possible design path is shown in 5, with the corresponding information in Figure 6. We can now predict whether users would be able to answer the sample questions efficiently when using a map with these assignments. The first question requires searching the symbol labels serially, because it uses the `Title` attribute, which is not mapped to a searchable property in this design. The remaining questions involve subsets of attributes on the design path, and are predicted to require quick visual search.

4.1. Mapping Mutual Funds

We illustrate the use of the flowchart with another example (Figure 7), in this case to design information maps of mutual funds. Suppose the available attributes are `Name`, `InvestmentObjective`, `YTDReturns`, `5YearReturns`, `Beta`, and `NetAssets`.

The first design (Figure 7, top) allows the user to weigh the tradeoff between volatility and performance for funds of

Question	**Query**
Was "Aliens" any good?	`Title = 'Aliens'`
Which recent comedies are well-reviewed?	`Year > 1995` **AND** `Genre = Comedy` **AND** `Review > 7.5`
When have the Oscars made an unpopular choice?	`Review < 7.5` **AND** `Oscar = Y`
What movies received Oscars in the eighties?	`1980 <= Year < 1990` **AND** `Oscar = Y`
Has any horror film ever received an Oscar?	`Genre = Horror` **AND** `Oscar = Y`

Table 2. Sample questions a user might use a Movie Map to answer.

a given size and investment objective. It uses a partition of the X axis to indicate the fund's investment objective, and a partition of the Y axis to indicate its size. Within each partition, the horizontal and vertical position of the symbol indicate its year to date and five year returns, respectively. The shape indicates the fund's volatility. In this way, five aspects of the fund can be concisely conveyed and searched.

A second design (Figure 7, bottom) uses the performance measures as the dominant selection criteria. The user can optionally animate a subset of funds by investment objective. The use of the flowchart thus facilitates the exploration of the map design space, suggesting multiple views that can accommodate different sets of query goals.

5. Related Work

Pirolli et al. have performed a study that tracks the eye movements of users as they perform textual navigation in a hyperbolic tree [20]. Their analysis finds that the size of the user's attentional spotlight is proportional to the local relevance of the text, and inversely proportional to the text's density. These results are relevant to creating information maps with large data sets, where views that let the user focus on a subset of information while retaining its context will be needed to prevent unusable information densities. Also, their theory focuses on how semantic (textual) cues serially guide the attentional spotlight, while this work focuses on perceptual, preattentive visual properties that the viewer can process in parallel. Clearly these processes are complementary and research to integrate such models is needed.

Card and MacKinlay present a formal way of describing the mapping of information types to visual properties [4]. Their goal is to analyze the structure of the larger information visualization design space, which includes trees, graphs, and hyperbolic views. Their goal is analysis, not design, but their representation would be useful if our work were extended to describe the perceptual constraints in the design of other types of information spaces.

Finally, the SAGE and VQE systems are part of a project to create a presentation design expert system [13, 6]. SAGE creates a media-independent design plan that fulfills the presentation's communicative goals. The design plan is given to a media allocator that can generate textual or graphic realizations of the plan. These design suggestions could be integrated into SAGE to assist its generation of information graphics. VQE adds *threads* that connect multiple views of the same data. Our proposal can be extended to accommodate multiple views by ensuring consistent visual semantics (e.g., never change the meaning of symbol colors across views).

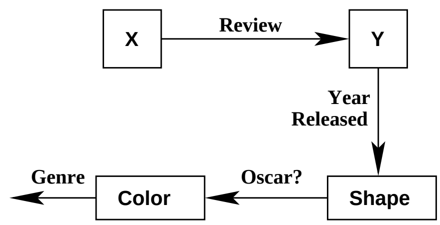

Figure 5. A design for the Movie Map, which is an example path in the flowchart in Figure 4.

6. Conclusion

This report explores the relationships among spatial selection, visual search, and visualization design. It proposes a method for designing information maps that makes use of experimental results regarding which kinds of visual search are efficient for viewers. These results were obtained from Wolfe's review of the experimental literature on visual search. The goal is to enable the map's users to do *query by attention* – answer questions by controlling visual attention (and receiving immediate feedback), rather than by manipulating the interface. Visual search for information is preferred because it is better suited for people, while database search remains better suited for computers.

While the application of experimental vision science to information design can result in oversimplification of both fields, we believe that useful design principles can be gleaned. Traditional textbooks on information design proceed by examining and critiquing design examples [26], and do not explicitly refer to results from vision science.

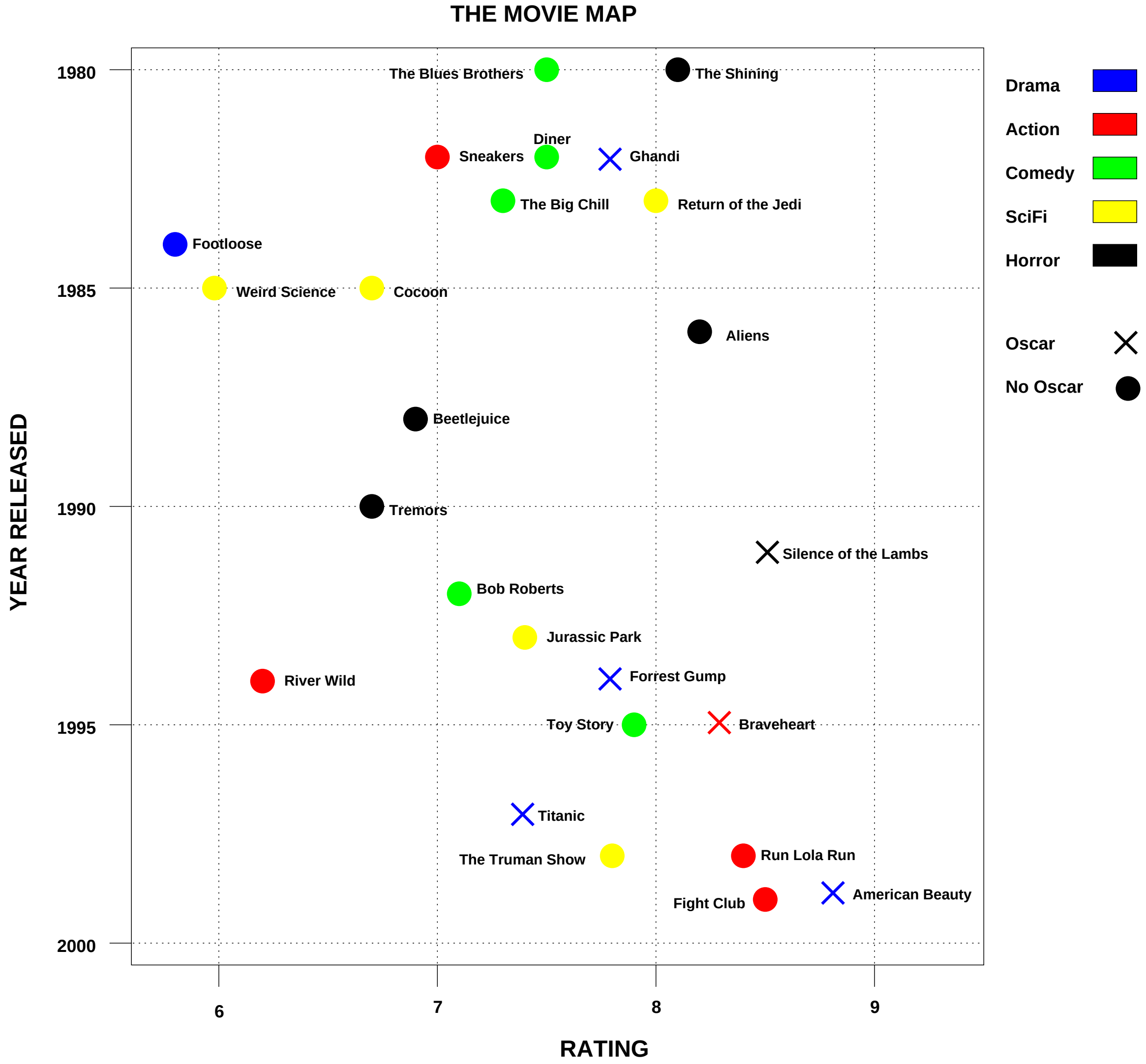

Figure 6. The Movie Map produced from the design path in Figure 5.

It remains to be seen whether this method can scale to real-world design problems, as the number of items in the map (100? 1,000?) and the number of attributes displayed increases. Techniques that give the user the ability to dynamically filter the data by manipulating sliders and toggles have proven effective for exploring large data sets [11, 2, 24]. And, dimension reduction techniques such as principal components analysis can simplify design (but only if the reduced dimensions are informative to users).

Further investigation of our visual capabilities can lead to new design insights. It remains to be seen whether people can search efficiently using arbitrary subsets of visual properties, so that more properties (such as texture, color, and shape) can be used at once. Also, most visual search experiments present a single type of target and distractor; understanding the role of heterogeneous distractors will help determine which visual properties can take more than two values. Using the heuristics discussed here in more design examples would further refine them, along with usability studies that can point out their shortcomings in practice.

It would also be worthwhile to see how well an empirically-based model of search performance could predict the time needed to complete map usage tasks. However, visual search experiments are conducted in carefully controlled conditions that may differ greatly from those in the day-to-day use of an information visualization. Users also tend to reformulate their information needs continuously. These factors would need to be taken into account in such an evaluation.

A future plan is to implement this proposal with an in-

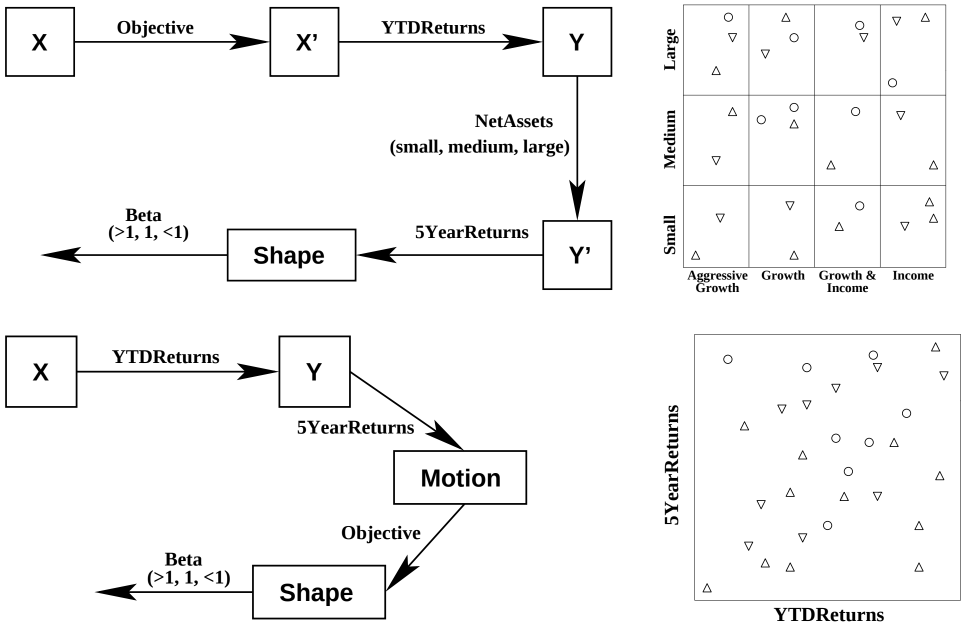

Figure 7. Two designs of maps for mutual funds. In cases where a real attribute is used on an edge marked E in the flowchart, the value ranges are shown in parenthesis. The figures at right suggest the appearance of the maps.

teractive tool that allows a user to rapidly prototype an information map given a small database or spreadsheet. The tool would let the user adjust the mapping of attributes to visual properties, while applying the kinds of perceptual constraints on efficient search presented here. (The tool's function would be similar to VQE [6].)

Efficient visual search in information maps is not possible in all cases. But maps that better employ our visual perception potentially have a great advantage over textual interfaces for understanding and managing information [5, 23]. I believe that visual information tools can benefit from progress in vision science, by offering design guidance related to our perceptual capabilities.

6.1. Acknowledgments

We thank Michael Oltmans, Tracy Hammond, Christine Alvarado, Glenn Evans, and Howie Shrobe (all at the MIT AI Lab) for helpful comments on an earlier draft. Thanks to Bart Anderson at the MIT Brain and Cognitive Sciences Department for his Visual Perception seminar that inspired this work. I also thank the anonymous reviewers for helpful feedback. This research is sponsored by the The MIT Ford Motor Company Collaboration and MIT Project Oxygen.

References

[1] C. Ahlberg and B. Shneiderman. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In *Proceedings of CHI '94: Human Factors in Computing Systems*, pages 313–317, Boston, Massachusetts, 1994.

[2] C. Alhberg and E. Wistrand. IVEE: An information visualization & exploration environment. In N. Gershon and S. G. Eick, editors, *Proceedings of the IEEE Symposium on Information Visualization (InfoVis '95)*, pages 66–73, Atlanta, Georgia, October 1995.

[3] J. Bertin. *Semiology of Graphics: Diagrams, Networks, Maps [orig. Semiologie Graphique]*. University of Wisconsin Press, Madison, WI, 1983. Translated by William J. Berg.

[4] S. K. Card and J. MacKinlay. The structure of the information visualization design space. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis '97)*, pages 92–99, October 1997.

[5] S. K. Card, J. D. MacKinlay, and B. Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, San Francisco, 1999.

[6] M. Derthick, J. Kolojejchick, and S. F. Roth. An interactive visual query environment for exploring data. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST '97)*, pages 189–198, 1997.

[7] J. Driver. Motion coherence and conjunction search: Implications for guided search theory. *Perception and Psychophysics*, 51(1):79–85, 1992.

[8] J. Driver. Reversing visual search asymmetries with conjunctions of movement and orientation. *Journal of Exp. Psychology: Human Perception and Performance*, 18(1):22–33, 1992.

[9] M. D'Zmura. Color in visual search. *Vision Research*, 30(6):951–966, 1990.

[10] H. E. Egeth, R. A. Vizri, and H. Garbart. Searching for conjunctively defined targets. *Journal of Exp. Psychology: Human Perception and Performance*, 10:32–39, 1984.

[11] S. Eick. Graphically displaying text. *Journal of Computational and Graphical Statistics*, 3(2):127–142, June 1994.

[12] J. B. Ellis, A. Rose, and C. Plaisant. Putting visualization to work: ProgramFinder for youth placement. In *Proceedings of CHI '97: Human Factors in Computing Systems*, pages 502–509, Atlanta, Georgia, 1997.

[13] N. Green, S. Kerpedjiev, and S. F. Roth. Generating visual arguments: a media-independent approach. In *AAAI98 Workshop on Representations for Multi-modal Human-Computer Interaction*, July 1998.

[14] H. Hocheiser. Browsers with changing parts: A catalog explorer for Philip Glass' Website. In *Proceedings of ACM Designing Interactive Systems Conference*, pages 105–115, Brooklyn, New York, 2000.

[15] P. McLeod, J. Driver, and J. Crisp. Visual search for conjunctions of movement and form is parallel. *Nature*, 332:154–155, 1988.

[16] P. McLeod, J. Driver, Z. Dienes, and J. Crisp. Filtering by movement in visual search. *Journal of Exp. Psychology: Human Perception and Performance*, 17(1):55–64, 1991.

[17] T. Mihalisin, J. Timlin, and J. Schwegler. Visualizing multivariate functions, data, and distributions. *IEEE Computer Graphics and Applications*, 11(3):28–35, 1991.

[18] K. Nakayama and G. H. Silverman. Serial and parallel processing of visual feature conjunctions. *Nature*, 320:264–265, 1986.

[19] S. E. Palmer. *Vision Science: Photons to Phenomenology*. The MIT Press, Cambridge, MA, 1999.

[20] P. Pirolli, S. K. Card, and M. M. V. D. Wege. Visual information foraging in a Focus+Context visualization. In *Proceedings of CHI 2001*, Seattle, Washington, 2001.

[21] R. Rosenholtz. A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, 39:3157–3163, 1999.

[22] H. S. Smallman and R. M. Boynton. Segregation of basic color in information displays. *Journal of the Optical Society of America A*, 7(10):1985–1994, 1990.

[23] R. Spence. *Information Visualization*. ACM Press/Addison Wesley, 2000.

[24] Spotfire. `http://www.spotfire.com/`.

[25] A. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136, 1980.

[26] E. R. Tufte. *Envisioning Information*. Graphics Press, Cheshire, Connecticut, 1990.

[27] J. M. Wolfe. Visual search. In H. Pashler, editor, *Attention*. University College London Press, London, 1996.