PREDICTIVE DYNAMIC USER INTERFACES FOR INTERACTIVE VISUAL SEARCH

Sam Mavandadi, Parham Aarabi, Azadeh Khaleghi, and Ron Appel

Artificial Perception Laboratory
University of Toronto

ABSTRACT

This paper proposes a method for designing user interfaces based on ideas rooted in data communication theory. It suggests that a visual user interface should be treated as a multi-transmitter, single-receiver communication system, where the total available bandwidth for transmission is limited. The proposed design entails the scaling of visual components that are displayed according to their degree of relevance to the user, or in other words, their probability of selection by the user.

1. INTRODUCTION

Visual computer interfaces are designed as gateways for users to access the raw information that is embedded within those systems. The design of these interfaces play key roles in the efficiency, speed, and ease with which a user can extract information. The efficiency of these interfaces has become more and more important over the past few decades with the prevalence of personal computers, personal digital assistants (PDAs), and mobile communication devices that are becoming smaller and smaller in size and visual display area year after year. In this paper, we propose that such an interface can be seen as a communication channel between a computer and a human. Furthermore, the overall system can be seen as a communication system with multiple transmitters and a single receiver. From this perspective, the concepts of information theory can be used to optimize the performance of such systems with respect to information rate and packing efficiency. We will show the implications to be that the amount of visualization area that is allocated to each particular object (image, application, etc.) must correspond to the user's level of interest in that object, which in turn corresponds to the probability of selection of the object by the user.

The concept of a variable visual representation size has been tackled in different contexts by researchers in the past [1, 2, 3, 4]. In [1], in the context of the *Collapse-to-zoom* methodology, the authors suggest that when displayed with more detail, the likelihood of users identifying more relevant content increases.

A fair amount of work has also been done with respect to resizing of icons to assist users in their selection [5, 6]. The conclusion was that although in cases where the number of objects to be selected is small these interfaces perform better, in situations where the selection is made among a large number of objects, the performance degrades.

In [7], the authors describe a space management scheme that dynamically allocates the space necessary for application windows that are opened on a computer screen. It offers a better management method over the conventional schemes that simply open the windows on top of others and allow for cascaded and tiled views on demand. To overcome these shortcomings, they offer a method that minimizes the amount of overlapping windows and at the same time attempts to maximize the amount of used area by each application.

In [8], the authors look at different methods of presenting photographs in order to assist users in the task of image searching. Different methods of presenting the photos were analyzed, including the arrangement of the photo thumbnails with respect to their similarity. It was determined that, on average, users prefer to have the images arranged such that similar photos are displayed close to one another.

In the following sections, we will describe the analogy between multi-transmitter, single-receiver communication systems and visual user interfaces. Then, we will develop a systematic approach to dynamically allocating the display area dedicated to objects based on their relevance to the user.

2. COMMUNICATION SYSTEMS

Consider the multiple access communication system shown in Figure 1, where each of the many transmitters independently transmits information on a common communication channel. Each transmitter i transmits a transmitter-dependent codeword of length C_i. The overall transmission is performed at discrete time intervals, with each individual transmitter independently deciding whether they would participate in each discrete transmission.

Since all the transmitters might be transmitting at the same time, the total system bandwidth B_T is limited and as a result, different transmitters, compete for bandwidth. Let us assume that each transmitter i is assigned a bandwidth B_i with the constraint that

$$\sum_i B_i = B_T \tag{1}$$

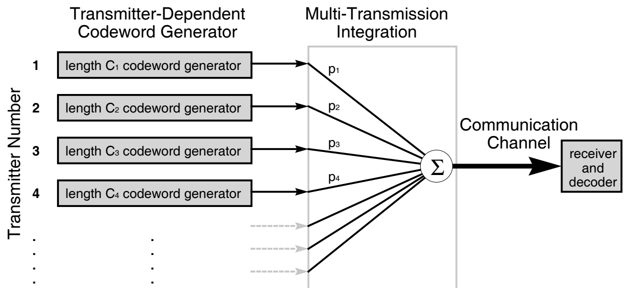

Fig. 1. A multiple access communication system, with multiple transmitters and a single common channel.

Now, the bandwidth occupied as a result of each transmission is inversely proportional to the codeword length. In other words, $C_i = \lambda B_i^{-1}$. This relationship arises as a result of the nature of frequency transforms, where the spectral bandwidth and the temporal length are inversely proportional to one another.

Now, for the different transmitters competing for the available bandwidth, an equal division (or, equivalently, equal codeword lengths) is often not the optimal strategy from an information theoretic perspective. The bandwidth B_i that is assigned to each transmitter is often scaled according to its transmission probability. For example, if in a dual-transmitter communication system the probability of A transmitting is higher than the probability of B transmitting, it makes sense to assign greater bandwidth for the transmission of A than the transmission of B.

Our goal is to choose appropriate C_i (or alternatively, B_i) values in order to minimize the average transmitted codeword length, while adhering to the total bandwidth constraint. The average codeword length L can be stated as:

$$L = \mathrm{E}\left[C_i\right] = \sum_i p_i C_i = \lambda \sum_i \frac{p_i}{B_i} \tag{2}$$

where p_i is the probability that transmitter i would transmit at each discrete time interval. It should be noted that since each transmitter decides independently when to transmit, there is no relationship between the transmission probabilities among the different broadcasters. It is easy to show using Lagrange multipliers that minimizing L subject to the condition $\sum_i B_i = B_T$ results in the following optimal bandwidth allocation:

$$B_i^* = \arg\min L = B_T \frac{\sqrt{p_i}}{\sum\limits_j \sqrt{p_j}} \tag{3}$$

where B_i^* is the optimal bandwidth that should be allocated to transmitter i.

3. DYNAMIC USER INTERFACES

We now consider a user interface where the user must select between a set of visual objects or images. The total area allocated to displaying these objects, A_T, is of course limited based on the size of the available display. The area assigned to image i is denoted as A_i with the following constraint:

$$\sum_i A_i = A_T \tag{4}$$

We propose that these image areas are in fact analogous to bandwidth in the communication system discussed in the previous section. Bandwidth and screen area are both limited resources, and a higher area dedicated to a particular item allows it to be identified more easily, more quickly, and to convey more information, just as a higher bandwidth allocated to a certain transmitter allows it to be communicate faster and to transmit more information. The optimization question in both cases is: how do we divide a limited resource to best convey information from multiple transmitters? In the one case, the transmitters are actual transmitters tied to a communication system, and in the other case, the transmitters are visual objects or images visually transmitting information to a human user.

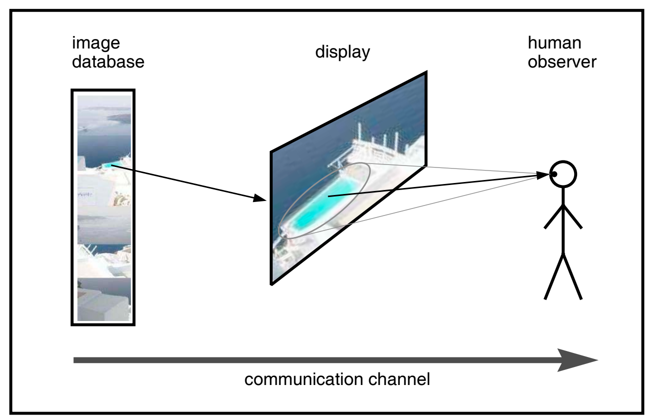

Fig. 2. A visual display depicted as a multi-transmitter communication channel.

As in the communication system example, we propose the following cost function to be minimized:

$$\Psi = \sum_i \frac{p_i}{A_i} \tag{5}$$

where p_i in this case is the probability that image i would be selected by the user, or alternatively, it is the probability that the user is seeking image i. Since a user may select or be interested in multiple images, the selection probability for each image is independent of the selection probabilities for the other images.

Our goal now is to minimize Ψ subject to the side constraint that $\sum_i A_i = A_T$, which is very much analogous to the bandwidth allocation problem considered previously. As in that case, it is easy to show that the optimal area assigned to image i is:

$$A_i^* = \arg\min \Psi = A_T \frac{\sqrt{p_i}}{\sum_j \sqrt{p_j}} \tag{6}$$

In other words, the area of each image/object being displayed should ideally be scaled with the square root of the probability that the user is interested in that specific image or object. The above discussion only pertains to the size of the displayed images, and not their location. Clearly, the location of the objects of interest is an important criterion. For example, consider the multiple images of Figure 3, consisting of a small database of images. If we assume that the object of interest is the light blue pool illustrated in Figure 4, then according to our previous discussion certain images that contain the pool should be scaled to be relatively larger than those that do not have a good view of this pool.

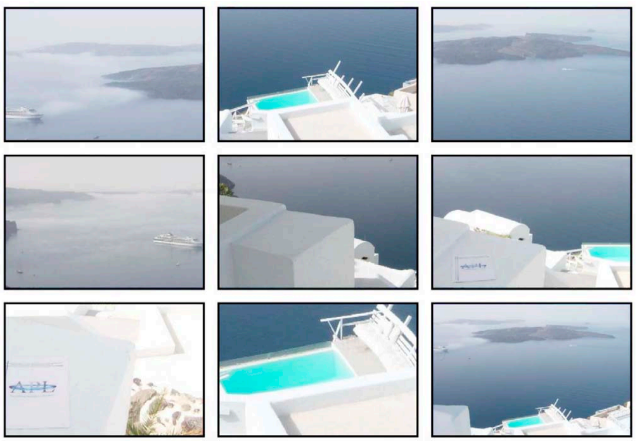

Fig. 3. A set of nine images of the island of Santorini in Greece.

A simple way of achieving this is by anchoring or centering the best image in the center, and horizontally surrounding each image with other smaller (and less likely to be chosen) images, as shown in Figure 5.

Another way is to anchor the primary image at the origin of a quadrant and surrounding the image within the quadrant with other smaller (and less likely to be chosen) images, as shown in Figure 6.

4. DESIGN EXAMPLE

In this section we will briefly outline the design of one such predictive interface. The problem that we are looking at is that of being able to quickly search and sift through large numbers of images that have been taken from the same environment. Large-scale surveillance systems are perfect examples of where the image database is conditioned in this manner.

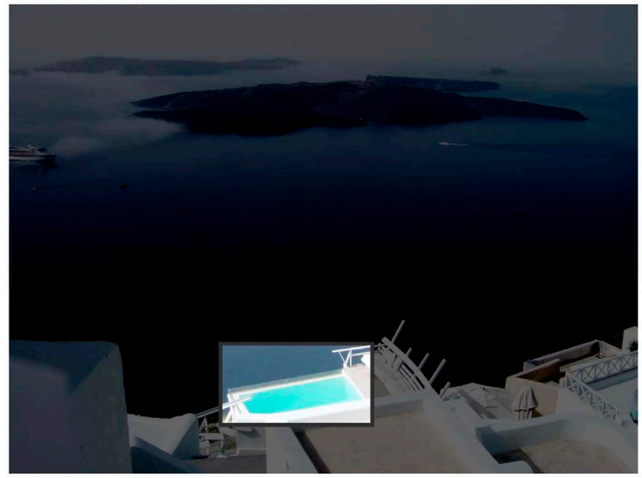

Fig. 4. A single object (the pool) is selected within a panoramic picture.

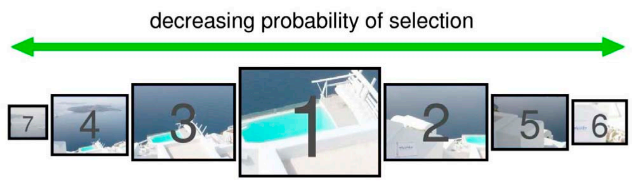

Fig. 5. One possible method of dynamically resizing the displayed images based on their relevance to the object of figure 4.

In the mentioned setting, there usually is more than one available vantage point from which the same spatial location can be observed from. Furthermore, a user may want to quickly find different observations of the same location from different vantage points. We propose that there should be an underlying interactive search mechanism that would attempt to find degrees of relevance among images, such that images that have with more overlapping contents would be more 'related' to one another. With this system in place, the user can select an image and the interface should display the most relevant images to the particular selection, scaled according to their degrees of relevance and allow the user to sift through them. This would be an alternative to the case where the user has to do an exhaustive search over all available images. The basic design of the proposed system is shown in Figure 7.

5. CONCLUSIONS

The design of a visual user interface is approached from a multi-transmitter communication system, where the human user is the lone receiver of the system. It was argued that the area that is dedicated to any visual component should be proportional to its probability of selection (or degree of rele-

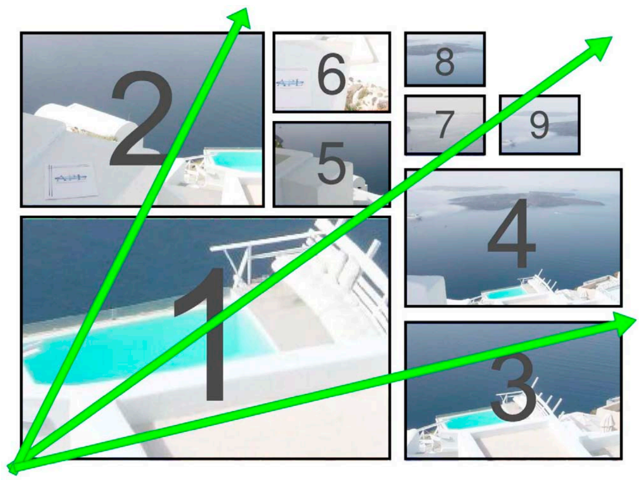

Fig. 6. Another possible method for dynamically resizing images based on their relevance.

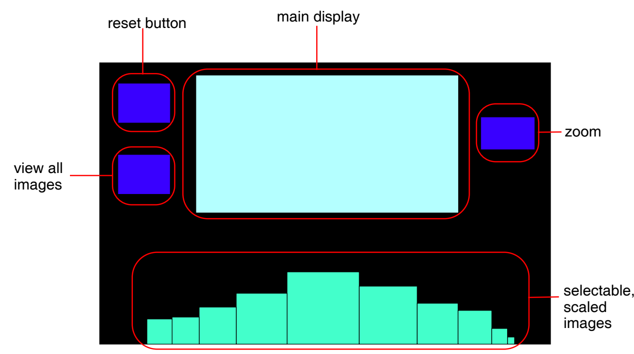

Fig. 7. Design example of a predictive dynamic interface. The current image is displayed in the centre of the image. Images relevant to the user's search are displayed on the bottom, and are scaled according to their relevance to the search. The user can interact with the search algorithm by selecting locations in the current image.

vance) by the user. This would allow the user to more quickly and efficiently find and select the object of interest. Furthermore, it would allow relevant objects to be displayed with more detail (i.e. higher visual resolution).

All of this brings us to the conclusion that for a visual user interface to be able to efficiently present the user with visual information, it also requires to do some pre-processing to determine what the user may be looking for. Therefore, we think that an interactive algorithm that can dynamically predict what the user is looking for should be a part of any visual interface technology.

6. REFERENCES

[1] Patrick Baudisch, Xing Xie, Chong Wang, and Wei-Ying Ma, "Collapse-to-zoom: viewing web pages on small screen devices by interactively removing irrelevant content," in *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, New York, NY, USA, 2004, pp. 91–94, ACM Press.

[2] Scott E. Hudson and Kenichiro Tanaka, "Providing visually rich resizable images for user interface components," in *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, New York, NY, USA, 2000, pp. 227–235, ACM Press.

[3] David J. Ward, Alan F. Blackwell, and David J. C. MacKay, "Dasher data entry interface using continuous gestures and language models," in *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, New York, NY, USA, 2000, pp. 129–137, ACM Press.

[4] John Lamping and Ramana Rao, "The hyperbolic browser: A focus + context technique for visualizing large hierarchies.," *Journal of Visual Languages and Computing*, vol. 7, pp. 33–55, 1996.

[5] Ravin Balakrishnan, "Beating Fitts' law: virtual enhancements for pointing facilitation," *International Journal of Human-Computer Studies*, vol. 61, pp. 857–874, Oct. 2004.

[6] Michael McGuffin and Ravin Balakrishnan, "Acquisition of expanding targets," in *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, 2002, pp. 57–64, ACM Press.

[7] Blaine A. Bell and Steven K. Feiner, "Dynamic space management for user interfaces," in *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, New York, NY, USA, 2000, pp. 239–248, ACM Press.

[8] Kerry Rodden, Wojciech Basalaj, David Sinclair, and Kenneth Wood, "Does organisation by similarity assist image browsing?," in *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, 2001, pp. 190–197, ACM Press.