Exploring Gradient-based Face Navigation Interfaces

Tzu-Pei Grace Chen
Department of Computer Science
University of British Columbia
tzupei@cs.ubc.ca

Sidney Fels
Department of Electrical and Computer Engineering
University of British Columbia
ssfels@ece.ubc.ca

Abstract

We have created a gradient-based face navigation interface that allows users to explore a large face space based on an eigenface technique. This approach to synthesizing faces contrasts with more typical techniques for forming composite faces based on the blending of facial features. We compare three ways of moving through the face space, using two types of sliders and a face-wheel. These are adapted from typical color space interfaces since they are commonly used. However, eigenface dimensions do not have meaningful text labels, unlike primary colors, necessitating the use of faces themselves for the labels of the navigation axes. Results suggest that users can navigate with face-labelled axes. They find slider interfaces best suited to finding the neighborhood of a target face, but that the face-wheel is better for refinement once inside the neighborhood.

Key words: Interaction techniques, information visualization, interactive computer graphics and perception and graphics.

1 Introduction

In Kyoto, Japan, there is a temple called Sanjusangendo that houses 1,000 life-sized Buddhist statues and one large central Buddha figure. Each statue has a unique face. The Japanese believe that it is possible to find all the faces of lost relatives among the sculptures. Inspired by this Japanese folklore, we look for new and better ways for people to find faces from a mental image in a large parameterized face space.

Faces are special. Gauthier et al. [10] pointed out that we recognize faces with finer categorization than objects and that we are expert at face recognition; thus, it is likely that specialized interfaces for faces are useful. Furthermore, face-related research is expanding but relatively little work is done on what specific user interface techniques support face navigation; therefore, it is an area of value to investigate.

We have created a face navigation system that allows users to explore a large face space based on an eigenface technique [2] for synthesizing faces in that space. This approach to synthesizing faces contrasts with more typical techniques for forming composite faces based on the blending of facial features. We compare three ways of moving through such face space, using two types of sliders (static and dynamic) and a face-wheel. They are adapted from typical color space interfaces. One significant difference is that dimensions in the principal component (eigenface) space do not have text labels that are meaningful to users, necessitating the use of faces themselves as the labels. We anticipate that showing face gradients along each dimension in the display will help users understand how changes in each dimension affect the face, and hence, allow them to find faces more easily. This approach requires a more sophisticated interface leading to the creation of our face-wheel interface, shown in Figure 2. Further details are explained in the "Design and Implementation" Section.

Our main research goal is to determine whether this interface is an effective way to navigate in this space compare to conventional slider interface. We hypothesize users' face recognition abilities will tune to the holistic differences in the faces displayed without the need for the explicit labels available with feature-based methods. We ran three pilot studies to narrow down the variables for our main experiment. The main experiment, described in the section titled "User Testing," compares a face-wheel interface to a slider-based interface for accuracy, speed and navigation patterns. Our main results indicate first that users can find target faces successfully using a holistic configuration approach even without labels. Second, we find that a slider-based interface is useful for finding the neighborhood of a target face, but the face-wheel is better for refining face navigation. Third, results indicate that, unlike colors, faces require further interface sophistication because of their higher dimensionality.

Unlike other face retrieval systems [23, 7, 3, 13], ours is novel because of several reasons. We focus on navigation aspect by having a structured face space in which to navigate. This is helpful in providing users better intuition and understanding. We avoided categorizing facial features as they have undesirable effects and can easily escalate to a large number. We also rely on color research in our interface design because it is of interest to make face creation and searching as easy as color selection. This is especially desirable where machines do not do a good enough job and users must rely on themselves to obtain a face close to their mental image.

Applications for face navigation interfaces include witness identification, creating new faces for the entertainment industry, computer animation, games and interactive artwork. Since technologies evolve around people, it is of interest to develop face navigation interface as it can facilitate face selection in applications. Moreover, our results have implications for the design of high-dimensional (less than 10) navigation interfaces in spaces where it is difficult to have meaningful labels.

The order of this paper is as follows: two, related work; three, the navigation systems; four, the experiments; five, discussion of the results and conclusion.

2 Previous Work

There are three main alternative methods for retrieving faces that are related to face navigation. These are: feature library retrieval, parametric retrieval, and statistical retrieval.

2.1 Feature Library Retrieval

Feature based retrieval typically provides a library of features for piecing together a face composite. There are several drawbacks to this method.

First, the retrieval relies on finding distinctive features; this works effectively for distinctive faces but less so with "typical" faces. Second, it encourages verbal categorization of facial features because face distinctiveness is more conveniently encoded in facial components [20]. This can lead to misidentification as Schooler et al. [16] have shown that verbal categorization affects memory; hence this is not favorable in the context of eyewitness testimony. Further, conventional face composite technologies that use the component methodology, such as *F*aces, *I*denti-Kit [13], *S*potIt! [3], *FACETTE* and *I*magine face the challenge of classifying facial features in their feature libraries. The possible categories can grow to be very large, even infinite [21].

Our approach avoids feature labelling and does away with specific features by focusing on configuring the whole face. This lets the user's ability to recognize faces and their differences guide navigation. This is desirable because psychological research supports human perceptual processing of face patterns to use holistic analysis rather than decompositions of facial features [11].

2.2 Parametric Retrieval

Another common approach is the use of humanoid faces to simplify the multiplicity of human faces. With prior knowledge of the parameter values used to generate and interpolate humanoid faces, face retrieval can be simplified to matching parameters. For face navigation, DeCarlo et al. [6] and DiPaola [7] have built systems for creating non-photorealistic humanoid faces. DeCarlo et al. use anthropometric statistics to generate faces through the use of variational modelling. DiPaola, on the other hand, uses a genetic algorithm to allow users to select "genotypes" that determine choices for a number of configural and component facial categories. This approach is like EvoFIT [12]. However, DiPaola's application is designed for a game, The Sims. Hence, it has a simple interface that allows users to browse easily through randomly generated faces and modify their looks through a number of parameter sliders. This method is generally used for entertainment avatars, as faces do not have to look real.

Our approach is also parametric; however, our faces look realistic and require a larger parameter space. Having realistic-looking faces is essential for our study as our application involves users' perception. Our navigation dimensions are based on eigenfaces computed by methods of Blanz and Vetter [2], suggesting that an image-based gradient can be created.

2.3 Statistical Retrieval

Recent developments in face retrieval have taken on a statistical approach. The use of principal component analysis with facial images has resulted in eigenface-based approaches [19] which, in turn, lend themselves to image-based retrieval systems such as *P*hotobook [15], *S*potIt! [3] and *C*AFIIR [23]. Baker [1] provides a good summary to these systems. In addition, eigenface methods can be applied to face reconstruction [2, 18] and evolution as in EvoFIT [12] as a way to retrieve faces.

While these techniques are useful for machine-based retrieval, we use the eigenface-based decomposition of face space [2] to define our face dimensions. This method allows faces to be linearly combined and separates texture and shape for better image coding and synthesis [22]. Thus, our face system anticipates users being able to navigate successfully for a Euclidean face space is easier to grasp, even though they do not represent specific features.

3 Design and Implementation

A block diagram of the system architecture is illustrated in Figure 1. We use FaceGen SDK (version 2.1) [8] to generate our face space. The programs are coded in C++, OpenGL and Tcl/Tk on the Windows 2000 OS. The navigation system displays the selected faces in an OpenGL window while its external control panel, built in Tcl/Tk, provides different settings. The two processes communicate through a socket.

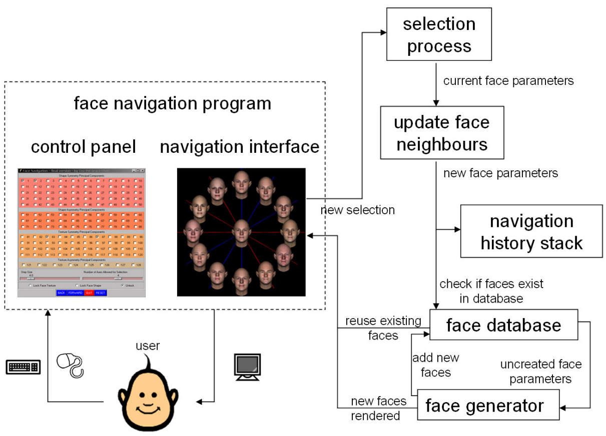

Figure 1: Block diagram of system architecture.

Our system can be divided into two main components: the face space in which users explore and the navigation interface.

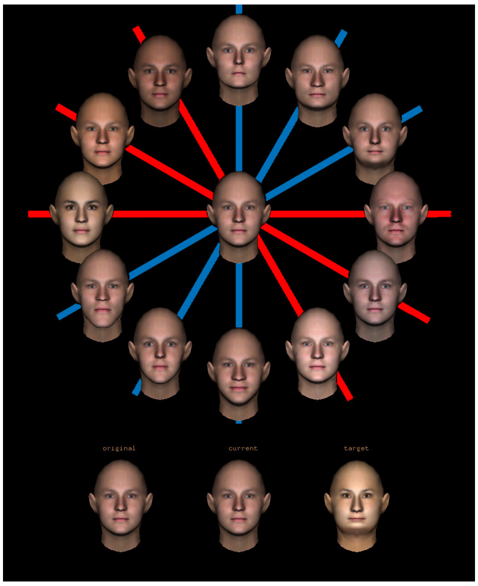

Figure 2: The wheel interface (coarse resolution) used for the experiment. The face line-up at the bottom is labeled in the order of "original," "current" and "target." The red lines indicate shape axes and the blue lines indicates texture axes. (Lines are thickened for clarity).

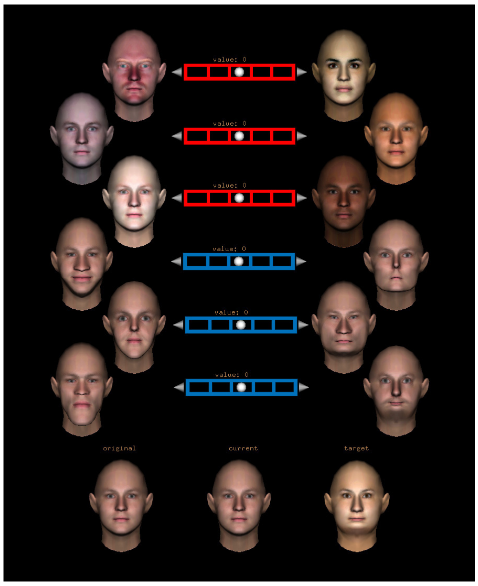

Figure 3: The slider interface (coarse resolution) used for the experiment. Just like the wheel interface (see Figure 2, it has a face line-up.

3.1 The Face Space

The well-established concept of face space, devised by Valentine [20], explains the various effects associated with face recognition. Although this face space model does not coincide with human perception, as pointed out by Busey [4], it is easy to understand and has gained strong support in psychology. The work by Blanz and Vetter [2] is a close interpretation of Valentine's norm-based face space model and the product FaceGen [8] essentially commercialized the work of Blanz et al. [2].

FaceGen provides a total of 128 principal components computed from 300 face scans. Because it uses aligned registered faces [2, 22], texture and shape principal components are obtained. There are 80 shape modes and 48 texture modes (of the shape, 30 are asymmetric, and of the texture, 8 are asymmetric). The 128 principal components of FaceGen serve as the basis vectors for FaceGen face space. Each, being one standard deviation in magnitude, describes different qualities of the face and can be used as navigation axes. These navigation axes can be thought of as "dimensions of manipulation," as changing values along the principal components resulting in different combinations of facial qualities. In this way, we can apply different principal components (which do not need to be used separately) to create different navigation axes and generate faces for our system.

By using FaceGen to populate our space, we have chosen the meanings of position, direction and distance in our face space. In every position of this space, there is a face whose qualities are defined by its location. Direction in this face space is determined by the qualities captured by the principal component in that direction. Depending on which principal component (or mixture of principal components) is used as a navigation axis, users can determine what face prototypes to expect by navigating in a certain direction. Distance traveled in this face space is related to the origin of the face space. The further the user moves from the origin, the further away he or she moves from the average face, and the less the current face resembles the average face. Further from the origin faces also have more prominent features, including exaggeration to the point of caricature.

Discretization and Optimization

Given that users can only view a subset of the face space at a time and computing resources are limited, it is necessary to discretize the face space and optimize the face displays for face navigation.

For testing purpose, we discretize the face space as such: given n, the number of navigation axes, and m, the number of positions allowed on the navigation axes, our face space has m^n number of faces. The value of m is essentially the resolution as we fix the range of each axis. Our studies use two values of m, namely 5 (coarse) and 9 (fine) and two values of n, namely 3 and 6 axes.

To facilitate our experiments, we optimize performance by precomputing faces where possible (i.e. when the selected parameters result in a relatively small face space). In cases where there are too many faces to precompute, we use a face cache with a pre-fetching mechanism as we can anticipate what users may select when the step size is set.

3.2 Face Navigation Interfaces

Although the number of facial dimensions has not been fully quantified, our concept is similar to how color interfaces work. If showing a color spectrum helps users match colors, then it may be useful to apply a similar technique to face matching.

Many color representation systems exist [17] and most colors can be created by some combinations of primary colors. Faces, unfortunately, do not have similarly agreed upon “primary faces.” However, in our approach, we are essentially assuming that the principal components used for navigation behave like primary colors do.

Like Schwarz [17], we are concerned with determining which navigation interface works best and in what situations. To this end, we have developed three different approaches. The wheel interface is described first followed by the two slider interfaces. They are shown in figures 2 and 3 respectively with a face line-up at the bottom to facilitate the face matching task used for testing.

Face-wheel Interface

Like a sliding window rather than a pie menu, the face-wheel shows local faces along respective axes around a currently selected central face. The current face is formed by the values along each of the principle components (presented as navigation axis) at the intersection. Its neighbors are one step away (either positive or negative) along each respective axis. To navigate, users click on the neighboring face that looks closer to the face they want. The entire arrangement then shifts so that the selected face becomes the central one with its local neighbors (Figure 4). This is similar to image color adjustment tools such as the one in Adobe Photo Deluxe.

We only show the one step neighborhood, however, additional ones can be shown at the expense of screen real-estate. The more dimensions placed in the face-wheel, the less space there is for showing face. We expect that up to about 10 axes could be shown, however, it is beyond the scope of this research to establish this. Our face-space has 128 dimensions, thus, we use a simple dimension selection interface in the control panel to allow users to select which dimensions are used in the wheel as shown in the left side of Figure 1.

Although only one round of neighbors forms the wheel as shown on the right in Figure 1, users can adjust the step size, via the control panel, to change the resolution of the face-wheel in order to see faces closer or further away from the current face. An alternate approach is to include more rounds of intermediate neighboring faces as the radius of the wheel increases. This, however, may be less efficient as it requires additional screen space and rendering power.

Another approach to visualize our face space is that of Design Galleries [14]. It is a method to facilitate visualization by showing users a range of possible results. We have not incorporated Design Galleries principles at this time as its construction phase is computationally intensive. Nevertheless, it is of interest to utilize this method to find suitable dispersion and arrangement algorithms for faces.

In addition to the basic face-wheel interface described above, our control panel provides other functions to facilitate navigation. Users can select up to six of the 128 principal components to be navigation axes by checking the appropriate boxes. We include a ”back” function, similar to the web browser, so users can retrace their paths. There is a locking mechanism to fix either the texture or the shape of the face while navigating. A control slider (previously mentioned) allows the adjustment of the step sizes of all navigation axes.

The Static Sliders

For experimental comparison, our first slider interface is made of discretized slider controls along the same axes as the face-wheel interface. It is intended to mimic conventional slider controls, although it is discretized, as our software cannot yet update faces fast enough to allow for continuous behavior. Hence, the slider behaves more like radio buttons. There is a lack of meaningful text labels at the ends of the sliders; therefore, we show what faces would result from moving the slider to extreme values relative to the average face at the origin. These end-point faces, however, remain fixed during the navigation process, hence the static sliders (Figure 4).

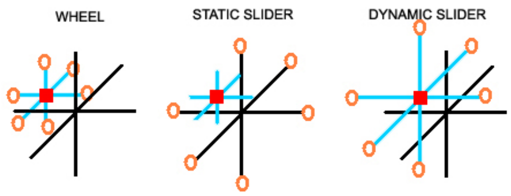

Figure 4: The figure illustrates conceptually the difference between the wheel, the static slider and the dynamic slider interfaces with three navigation axes. The solid red square indicates the current face. The orange ovals represent the face neighbors displayed. The black axes are the global axes of the face space and the light blue axes of the red square are the local axes of the current face.

The Dynamic Sliders

Static sliders have a clear disadvantage in that they only show extreme neighbors of the average face at every navigation step. Therefore, we provide dynamic sliders for a fairer comparison to the face-wheel in our experiments. These dynamic sliders are a hybrid of the wheel and static slider interfaces. The main attribute of dynamic sliders is that the end points are all updated each time the user moves (Figure 4). The reason this may be important is that changes on each axis are dependent on the values of the other sliders. If compared to equivalent RGB color sliders, it would be the same as showing the color at the end of each range of each slider, given the current slider settings. We suspect that this feature may be useful, since users can see the effect the extreme value on any axis has at every moment in time. This update gives users partic-

	Wheel	Static Sliders	Dynamic Sliders
Move-ments	One step at a time	Multiple steps	Multiple steps
Faces shown	Nearest neighbors of current	Extreme neighbors of origin	Extreme neighbors of current
Face update	Dynamic	Fixed	Dynamic
Extra	"back" function for undo	None	None

Table 1: Features of the navigation interfaces for testing.

ularly good visual cues if they reach the neighborhood of a target face.

To ensure the fairness of interface comparison, the wheel interface, for user testing, does not include the control panel step size adjustment slider. It is therefore limited in that it only allows users to move one step of a set size at a time. The slider interfaces, on the other hand, are not as restricted in the number of steps subjects can move at a time. A summary of the interface feature differences (as used in the experiments) is shown in Table 1.

4 User Testing

The experiment involves a face matching task where users are provided with explicit images to find in the database. We chose a matching task as this provides a way to control the image across subjects for study. In contrast, our target application is face-navigation where users have only a mental image of the face they want to find, however, we anticipate that using a known target image will provide much of the insight into the general task. We also hand select the range and principal components to make sure the faces stay within a realistic zone and have a significant amount of variance. We conducted three pilot studies [5] to narrow the conditions for our final experiment.

4.1 Pilot Studies

From our three pilot studies [5] we notice that three axes are too few to show performance variation. This is consistent with [17] for color. The one-to-one mapping of navigation axis to principal component versus one-to-two correlated mapping does not make much difference. The static sliders force subjects to exhibit a large trial and error pattern because the fixed face labels were not particularly helpful. The resolution appears to interact with interface types and users appear to use the various interfaces differently, depending upon how close to the target they are. Also, if resolution is too high, the wheel interface does not help much, as displayed faces become indistinguishable. Finally, the target location appears to have a counter-intuitive effect; that is, the further away the target is at the start, the better the performance. We suspet this is related to the distinctiveness of faces that are further from average face.

4.2 The Experiment

Given these findings, for our main experiment, we look at the ways users use the different interfaces as they approach the target more closely. We test three independent variables: interface (wheel and dynamic sliders), resolution (coarse and fine) and target type (near and far). The experiment is a **within-subject factorial** design, in which a total of 16 trials from 8 different conditions (two trial/condition) are randomly given to each subject to avoid order effects. Each trial is a face-matching task where subjects click on the interfaces to manipulate the face display. The face space in which they navigate is 5^6 for coarse resolution and 9^6 for fine resolution. With our factorial design experiment, we analyze the data using analysis of variance techniques.

Subjects

We recruit fifteen volunteers as subjects, of whom seven are male and eight are female. The group includes three Caucasians, two East Indians and nine Asians. None of the subjects has specialized training with face recognition.

Apparatus

The interfaces used in this experiment are the wheel and dynamic slider. The simplified wheel interface does not allow users to access all 128 principal components, hence the control panel is replaced by a simpler one. Both interfaces have six preset navigation axes and step size.

We choose face targets for each condition in a way that requires subjects to make the same amount of moves on the navigation axes. This ensures that targets of the same condition are spatially equidistant from the origin in the discrete face space. We ensure that the targets are all realistic looking and avoid caricatures.

We use six independent axes, of which three provide symmetric face shape variations and three provide symmetric face texture variations. We select these carefully from navigating the 128 principal components extensively during our pilot studies to make sure that changes are distinguishable along each dimension. This is consistent with Schwarz's finding [17], where subjects do not exhibit a large variation in color navigation performance with different mappings (whether via RGB or CMY), as long as the principal colors are quite distinguishable. Out of the subset that fits the criteria of large variations and realistic-looking facial features, we randomly picked six. Perceptual testing may provide finer selection criteria, but given that the high order principal components make clear and obvious differences to the faces, we leave that verification for future work. The axes and principal component mappings are shown in Table 2 with step sizes indicated for a coarse resolution. The range is selected to give significant variation over the values available for each axis, but not large enough to exaggerate faces to caricatures. The coarse resolution step sizes for the axes are twice as large as the ones for fine resolution, thus, twice as many steps are required to get to the same destination when navigating with the fine resolution setting.

Axes	PC	Range	Coarse Step Size
A_1	S_6	[-14; 14]	7
A_2	S_8	[-12; 12]	6
A_3	S_{10}	[-10; 10]	5
A_4	T_1	[-2; 2]	1
A_5	T_3	[-4; 4]	2
A_6	T_7	[-10; 10]	5

Table 2: Navigation axes used in the final experiment. Values for fine resolution step sizes are half of the corresponding coarse resolution step sizes. The units are in standard deviation as reported from FaceGen.

Procedure

As an example, a typical trial using the face-wheel proceeds as follows: 1. a new face target is shown in current interface condition; 2. subject clicks on face in the face-wheel that moves them closer to the target face; 3. subject continues until center face matches target face; 4. subject is given signal that face has been matched and new trial begins. Similarly, subjects use the sliders to try to match the target face in the slider interface conditions.

Each subject takes approximately 30 minutes to one hour for the test. Subjects also practice using each interface with different resolutions until comfortable with the interfaces. This generally takes 20-30 minutes. During practice, however, targets are randomized with varying difficulties. The target conditions set for the tests are deliberately excluded so that users cannot memorize paths during practice.

During testing, subjects' navigation patterns are recorded by registering the locations in face space visited and the time stamps of each move made. After every trial, users are instructed by a message above the face line-up to proceed to the next trial. There is no time-restriction enforced on the trial; only a reminder that if a trial takes more than five minutes, subjects can choose to abandon the trial and proceed to the next. After the test, subjects are interviewed for additional comments on the interfaces.

5 Results

The dependent measures are: time, accuracy score, repetitions (location revisits), refinement counts and approaching counts. ("Refinement" counts are the number of faces visited that fall within a two-step neighbor of the target; "approaching" counts are the number of faces visited outside of this scope. A two-step neighbor is chosen as that is where differences are observed.) We report our findings by effects of factors.

5.1 Resolution

As expected, resolution is significant with all dependent measures. With time (F(1,5)=88.86, MS_e=14189.80, p<0.01), the significance suggests that the finer the resolution, the more time subjects need to reach the targets. With accuracy scores (F(1,5)=16.87, MS_e=0.08, p<0.01), the significance reveals that the finer the resolution, the lower the subjects score. With repetition counts (F(1,5)=28.72, MS_e=23.33, p<0.01), significance suggests that the finer the resolution, the more repetitions there are. This is further echoed in the significant interaction between the resolution and the interface factor (F(1,5)=7.08, MS_e=18.43, p<0.05). Follow-up main effect analysis reveals that interaction is only significant (F(1,5)=13.36, p<0.05) with fine resolution and wheel interface. With refinement counts (F(1,5)=34.18, MS_e=25.24, p<0.01), the significance indicates that the finer the resolution, the more faces subjects sift through in order to reach the goal. As for approaching counts (F(1,5)=34.96, MS_e=112.30, p<0.01), the finer the resolution, the more steps are required to get into close proximity.

5.2 Interface

The interface is only significant when dependent measures are repetition counts and approaching counts. With repetition counts (F(1,5)=5.28, MS_e=25.84, p<0.05), the significance suggests that subjects using the wheel interface have made more repetitions with fine resolution, as confirmed in the main effect analysis of the resolution interaction, described previously. As for approaching counts (F(1,5)=16.81, MS_e=54.05, p<0.01), the significance indicates that subjects have to step through more face locations using the wheel interface than they do using the slider interface, before reaching the two-step neighborhood. As mentioned, this is expected as the wheel interface can only move one step at a time.

5.3 Target Position

The target position is only significant when the dependent measures are time and approaching counts. With time (F(1,5)=8.21, MS_e=15189.90, p<0.05), the significance indicates that the further away the target, the more time is required. With approaching count (F(1,5)=24.04, MS_e=63.86, p<0.01), the significance implies that the further away the targets, the more steps subjects take to reach the neighborhood.

6 Discussion

6.1 Resolution

We expect resolution to be an obvious source of variation and our ANOVA result indicate that consistently. We observe that subjects find the wheel interface to be particularly difficult to use with fine resolution. This is probably due to the fact that face prominence is below the just-noticeable-difference (JND) threshold. Thus, having gradient information at local regions is not very helpful at high resolutions. Dynamic sliders, on the other hand, overcome this problem by showing extreme faces. People can estimate the face gradients between extreme faces, although this comes by way of trial-and-error; that is, by guessing the face gradients in between. This suggests that it is very helpful to provide the global direction of the face space so that users can recognize, rather than recall, where to go.

6.2 Interface

We observe that subjects appear to visit fewer faces when using the wheel interface during refinement as compared to dynamic sliders. On the other hand, they approach the neighborhood more efficiently when using the dynamic slider interface. Figures 5 and 6 show the comparison. This indicates that perhaps people do not need the local gradient information at the beginning of navigation. All they need to do is to get into the correct zone by using sliders. Once they are close enough, local gradient information can help with refinement and thus help subjects reach the target. However, if the resolution is too fine, then it is useful to implement mechanisms that allow side-by-side comparisons. This trend implies that, in order to minimize the number of visits, sliders may be better for reaching the neighborhood and the wheel interface may be better for refinement.

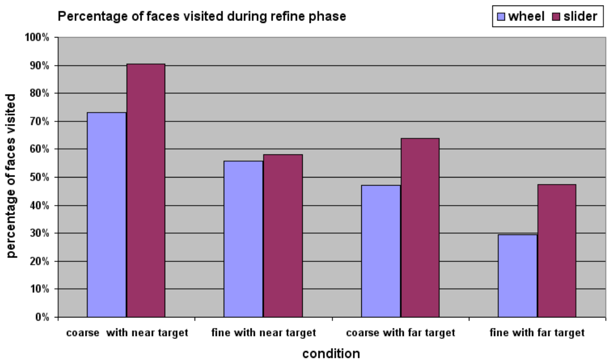

Figure 5: The histogram shows a higher percentage of faces visited via sliders during the refinement stage. The condition labels are "coarse with near target," "fine with near target," "coarse with far target" and "fine with far target" from left to right.

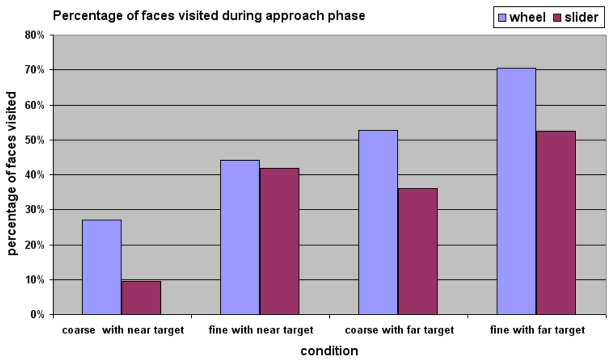

Figure 6: The histogram shows a higher percentage of faces visited via the wheel interface during the approaching stage.

6.3 Target Position

The effect of target position is what can be expected. Although we suspected, from our pilot studies, that targets further away are more distinguishable and hence may be easier to retrieve, this trend is not observed in our main experiment.

7 Conclusion and Future Work

In this paper, we describe the design and implementation of a face navigation system, based on configuring the whole face rather than selected features. The usage is to facilitate face browsing and selection in face related applications. Since face space navigation interface is a relatively new field, we relied on more defined work, the color space user interface, to define our approach.

Given the high dimensionality of faces, numerous static sliders together can be hard to use. The wheel interface, although restricted in its local view, provides a preview of possible combinations for selection. A hybrid system, such as dynamic sliders, allows faster global access but is still less effective than the wheel interface for refinement. Although quantitatively, the wheel and the dynamic slider interfaces do not differ much in performance, qualitatively, they lead to different behaviors of subjects. This illustrates that there are benefits in providing face gradients and that more sophisticated face interfaces are useful.

Although our system is not ready to provide the full effect of Sanjusangen-do where visitors can expect to find their lost relatives, we have shown that a gradient-based mechanism is suitable for refinement in face matching tasks. In addition, it is also useful for tackling the high dimensionality of faces since it avoids verbal categorization of facial features. Our system is limited in its lack of a systematic way to present all 128 face prototype axes. To do so may require categorizing faces in some perceptually sensible way and designing a face similarity metric, which is outside the scope of our research at this time.

For future work, we want to increase the number of navigation axes and rounds of neighbors to see how it would optimally balance screen real estate, usability and run-time performance. We plan to investigate the effects of other parameters, such as the starting position of navigation and the JND thresholds of users. Furthermore, a larger face space will allow for a qualitative study of finding faces that are not explicitly represented by photos and may or may not exist in the database. Improving the speed of face construction will allow for continuous navigation axes, providing additional aspects to investigate. It may also be useful to try to model the face navigation time by a similar equation to Fitts' Law [9].

As face synthesis technology improves and applications for witness identification and face creation in entertainment, computer animation and games become more important, having an easy-to-navigate face space becomes critical. We have established the benefits and constraints in the navigation system built. Given that its mechanism is independent from the face generator, we are optimistic that this gradient-based method has potential applications in complementing existing face retrieval systems and can extend to the visualization and navigation of higher dimensional spaces.

Acknowledgements

We would like to thank Singular Inversions, the members of the HCT Lab, Maria Klawe, Brian Fisher, and our subjects for their contributions. Funding for this research has been provided by ATR Media Information Sciences (MIS), Japan. We would also like to thank the anonymous reviewers for their helpful comments.

References

[1] E. Baker. *The Mug-Shot Search Problem - A Study of the Eigenface Metric, Search Strategies, and Interfaces in a System for Searching Facial Image Data*. PhD thesis, The Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, January 1999.

[2] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *Proc. of SIGGRAPH*, pages 187–194, 1999.

[3] R. Brunelli and O. Mich. SpotIt! and interactive identikit system. In *Proc. of GMIP*, volume 58, pages 399–404, 1996.

[4] T. A. Busey. Physical and psychological representations of faces: Evidence from morphing. *Psychological Science*, 9:476–483, 1998.

[5] T. P. Chen. 1,001,001 faces: A configural face navigation interface. Master's thesis, Department of Computer Science, University of British Columbia, Vancouver, British Columbia, June 2003.

[6] D. DeCarlo, D. Metaxas, and M. Stone. An anthropometric face model using variational techniques. In *Proc. of SIGGRAPH*, pages 67–74, 1998.

[7] S. Dipaola. Facespace:a facial spatial-domain toolkit. In *Proc. of IV*, pages 105–109, 2002.

[8] FaceGen. <`http://www.facegen.com`>.

[9] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movements. *Journal of Experimental Psychology*, 47:381–391, 1954.

[10] I. Gauthier, P. Skudlarski, J. C. Gore, and A. W. Anderson. Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3:191–197, 2000.

[11] P. J. B. Hancock, A. M. Burton, and V. Bruce. Face processing: Human perception and principal components analysis. *Memory & Cognition*, 24:26–40, 1996.

[12] P. J. B. Hancock and C. D. Frowd. Evolutionary generation of faces. In *Proc. of AISB*, pages 93–99, 1999.

[13] K. R. Laughery and R. H. Fowler. Sketch artist and Identi-Kit. procedure for recalling faces. *Journal of Applied Psychology*, 65(3):307–316, 1980.

[14] J. Marks, B. Andalman, P. A. Beardsley, W. Freeman, S. Gibson, J. Hodgins, T. Kang, B. Mirtich, H. Pfister, W. Ruml, K. Ryall, J. Seims, and S. Shieber. Design galleries: A general approach to setting parameters for computer graphics and animation. In *Proc. of SIGGRAPH*, pages 389–400, 1997.

[15] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Tools for content-based manipulation of image databases. In *Proc. of SPIE Storage and Retrieval Image and Video Databases II*, volume 2185, 1994.

[16] J. W. Schooler and T. Y. Engstler-Schooler. Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22(1):36–71, 1990.

[17] M. W. Schwarz, W. B. Cowan, and J. C. Beatty. An experimental comparison of RGB, YIQ, LAB, HSV and opponent color models. *ACM Transactions on Graphics*, 6(2):123–158, 1987.

[18] C. Tredoux, Y. Rosenthal, L. da Costa, and D. Nuenz. Face reconstruction using a configural, eigenface-based composite system. In *Proc. of SARMAC III*, 1999.

[19] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proc. of CVPR*, pages 586–591, 1991.

[20] T. Valentine. A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology A*, 43:161–204, 1991.

[21] T. Valentine. Face-space models of face recognition. In M. J. Wenger and J. T. Townsend, editors, *Computational, geometric, and process perspectives on facial recognition: Contexts and challenges*, pages 83–113. Lawrence Erlbaum Associates, Mahwah, New Jersey, 2001.

[22] T. Vetter and N. F. Troje. *Journal of the Optical Society of America A*, 14(9):2152–2161, 1997.

[23] J.K. Wu, Y.H.Ang, P. Lam, H.H. Loh, and A.D. Narasimhalu. Inference and retrieval of facial images. *Multimedia Systems*, 2(1):1–14, 1994.