Perception of Parameter Variations in Linear Fractal Images

Daryl H. Hepting and Leila Latifi

Computer Science Department, University of Regina, Canada
dhh@cs.uregina.ca,latifi2l@uregina.ca

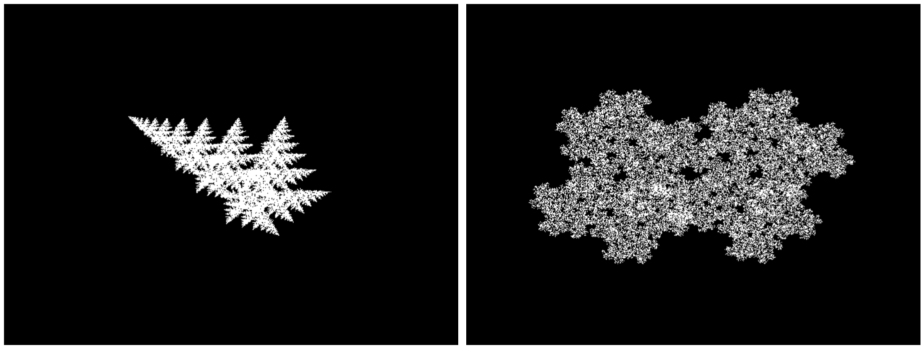
Abstract. Parametric images, defined by a small number of parameters, may help to democratize access to image creation because simple parameter manipulations can yield interesting variations. Foe example, many people appreciate the aesthetics of fractal images, but few are inclined to engage in the mathematics needed to create them. A perception-driven interface for fractal image creation could find a wide audience as people could use it as an outlet for their own creative expression. This paper discusses some first steps along that path, with a study and analysis of how participants perceived changes between smoothly varying images. Further steps towards a perception-driven interface are then laid out.

1 Introduction

The study of human visual perception has been addressed by many researchers who have, for example, studied humans' reactions when they see different images. When these images are computer-generated fractals, it is possible to gain additional information because the parameters used to generate each image are known. Therefore, they seem to be ideally-suited as stimuli for a perceptual study.

Linear fractal images are generated by iteratively applying a set of linear transformations (each comprising rotations, translations and scalings) to an initial point. These images have the very desirable property of "database amplification" [1], meaning that the images can be encoded by very simple and compact equations. Furthermore, while images close together in parameter space are similar, over the range of parameters there can be stunning differences between generated images. Two samples of images used in the perceptual study are shown in Figure 1. Although fractals have been studied in many ways, this study of how they are perceived, and the connection between that perception and the input parameters used to generate the images, is new.

Can an individual's exploration of this image parameter space (here, a two dimensional space defined by 2 rotation angle parameters, θ_1 and θ_2) be augmented with support to more easily find and understand surprising images? In his book, *Digital Harmony*, Whitney [2] talked about exploring for months within the parameter space he chose for "Arabesque" and not finishing. Can technology help the artist in this case? If we can understand how people perceive these images in



(a) (b)

Fig. 1. Two examples of the linear fractals used in the study. On the left, the values for the rotation angle parameters, θ_1 and θ_2, are 0.0 and 18.0 degrees. On the right, these values are 28.0 and 31.0 degrees.

relation to one another, we have the possibility to create a much more intuitive interface for image creation [3]. In turn, this can empower many more people to express their creativity, which could lead to more user-generated content on video sharing sites like youtube (http://www.youtube.com).

The rest of this paper is organized as follows. Section 2 provides a review of relevant literature. Section 3 introduces the approach that was developed to generate a manageable stimulus set. Section 4 presents the experiment's results and the method used to evaluate and analysis the results. To develop a baseline for the perceptual study, the Section 5 presents a discussion of the results along with directions for future work.

2 Background

Fractals are simple and recursive geometric objects that can be divided into many parts, and in which each part is a smaller, possibly rotated, copy of the whole. They can be magnified an unlimited number of times and their structure will be the same in every magnification. Benoit Mandelbrot [4] coined the term fractal – "to break" or "to fragment" which comes from the Latin verb "frangeres" – although he was not the first to study them.

Julia sets [4], defined by a complex quadratic equation, are some of the most famous fractals. The Julia set corresponding to $c = -1$ is shown in Figure 2. Considering the real and imaginary parts of the complex number, two parameters may be manipulated.

Linear fractals (see Figure 1) are conceptually similar to quadratic fractals, but permit a wider range of parameter manipulations. They are defined by a set of contractive linear transformations. Each of these linear transformations can be described in terms of rotation, scaling, and translation. Scaling and Translation each have two parameters, one for x and one for y while Rotation just

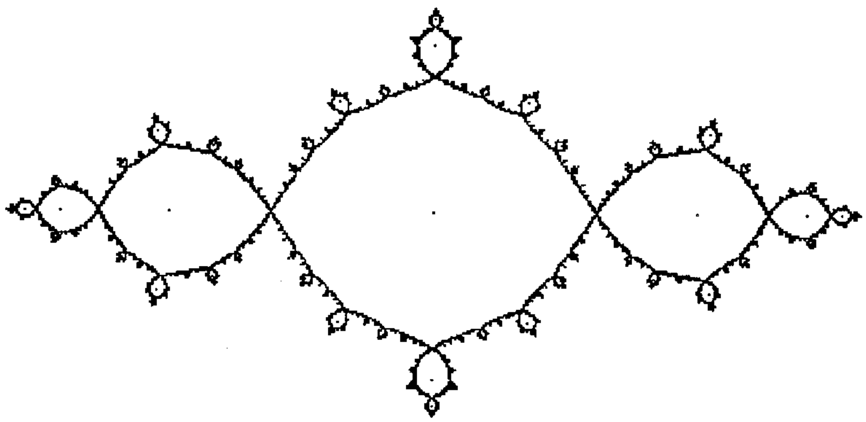


Fig. 2. An example of a Julia set, for $c = -1$.

has one parameter, θ. In total, five parameters may be manipulated for each transformation defining a linear fractal. Intricate shapes arise when two or more transformations are used.

To better understand how people perceive fractal images, we examine both computational and perceptual perspectives, to see how they can inform each other.

2.1 Computed Metrics

In order to provide a baseline for human judgements, we computed various metrics from the images.

Dimension Fractals are generally quantified by a non-integer dimension, which indicates how completely the fractal fills the space. The examples from Figure 1 are not lines but neither do they fill the plane. To approximate fractal dimension, we use a box-counting procedure [5]. As the size of box goes to 0, we count how many boxes are required to cover the fractal, and determine the dimension according to the following equation (where s is the box size):

$$D = \lim_{s \to 0} \frac{\log(N(s))}{\log(1/s)}. \tag{1}$$

Lacunarity Two fractal patterns could have the same dimension but look different. Whereas fractal dimension is a measure of how much space is filled by a fractal, lacunarity is a complementary measure of how it fills the space [6]. The gliding box algorithm [7] is a very popular method to calculate lacunarity. According to this algorithm a box of size r slides over an image, and the number of pixels inside that box is counted to determine the mass, M. $n(M, r)$ is the number of a gliding-boxes with radius r and mass M. $Q(M, r)$ is the probability distribution obtained when $n(M, r)$ is divided by the total number of boxes of size r. Lacunarity at scale r, denoted $L(r)$, is defined as the mean-square deviation of the variation of mass distribution probability $Q(M, r)$ divided by its square mean.

$$L(r) = \frac{\sum_M M^2 Q(M, r)}{[\sum_M M Q(M, r)]^2} \tag{2}$$

Mean Squared Error Mean Squared Error (MSE) is a pixel-based method for comparing two images, one of the most important problems in image processing. The definition of the MSE metric is given in Equation 3 where the images to be compared, I and K, are of size $m \times n$. This method is popular because it is simple, though it is not found to be a good match for human perception. It is used here to quantify the difference between successive images.

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} ||I(i, j) - K(i, j)||^2 \tag{3}$$

Principal Component Analysis Principal component analysis (PCA) is a standard method in exploratory data analysis to decrease the dimensionality of data by identifying patterns in data, highlighting the similarity and differences and extracting the meaningful variables [8]. To assess image similarity, we provide a training set of images from which vectors are computed. Standard practice indicates that we need only choose enough vectors (say N) to account for at least 80% of the variance. We then project all images onto these chosen vectors to then represent each image as a point within this N-dimensional space, from which it is easy to assess similarity based on proximity.

2.2 Perceptual Evaluation

Both the quantity and quality of stimuli are important factors to consider when designing a study. To make a thorough comparison of a set of stimuli, the participant should make judgments about all possible pairs of stimuli, which is simply not feasible in all but simple situations [9]. For this reason, much effort has been placed on developing games as a way to encourage the crowdsourcing of comparisons [10].

Our study employed the method of limits, which involves the presentation of ascending and descending series of stimuli varying along a single dimension. The difference threshold is the smallest amount of change in intensity that is yields a "just noticeable difference". In this study, the difference threshold is calculated as the change in a stimulus that is required to elicit a response of "different" from the participant. By doing this in both an ascending and descending series, two estimates of the difference threshold are obtained.

3 Experiment Design

In order to assess viewers' perceptions of fractal imagery, it is first important to select the set of stimuli. The fractals used in this study consist of two transformations, which each have: 1 rotation angle parameter, 2 scale parameters, and 2 translation parameters. If 10 values are allowed for each parameter, $10^{10} = 10,000,000,000$ images would be generated. It is important to select a subset which will not overwhelm any participant, yet be representative of what is possible from the parameter space. Barnsley [11] proved that small changes in input parameters results in small changes in the image, therefore it is worthwhile choose samples from the whole parameter space. Therefore the scaling and translation parameters were fixed and the rotation angles were sampled every 3.6 degrees in order to create a grid of $100 \times 100 = 10,000$ images.

3.1 Method

How should the study be structured in order to gain as much useful information as possible, and as accurately as possible? The 10,000 images in stimulus set were divided into incrementally-varying sequences of 100 images. Each participant was shown eight sequences, in both the forward and backward directions (as dictated by the method of limits) to see how he or she could detect changes related to the rotation angle parameter value changes.

A web-based program was implemented to run the study (see Figure 3 for a screenshot). The application randomly presented each participant with these sequences (4 from rows/columns and 4 from diagonals in the grid) in forward and reverse directions for a total of 16 sequences. Each sequence consisted of 100 images, which were shown two images at a time. Each image, aside from the first and last, was on the screen in the right position first, then the left. The participants' task was to click the "changed" button anytime they thought the two images on the screen were noticeably different.

We ran 2 conditions of the study, each with 25 participants drawn from the undergraduate and graduate student populations at the University of Regina. In the first condition, image sequences were drawn from the columns and diagonals of the image grid (see Figure 4, left) − 4 columns and 4 diagonals were randomly selected. In the second condition, image sequences were drawn from the rows and opposite diagonals of the image grid (see Figure 4, right) − 4 rows and 4 opposite diagonals were randomly selected. Each sequence was viewed in both forward and backward directions. Therefore every user saw 16 sequences of 100 images for a total of 1600 images per participant. In each condition, each image was rated as part of a column or a row and as part of a diagonal or an opposite

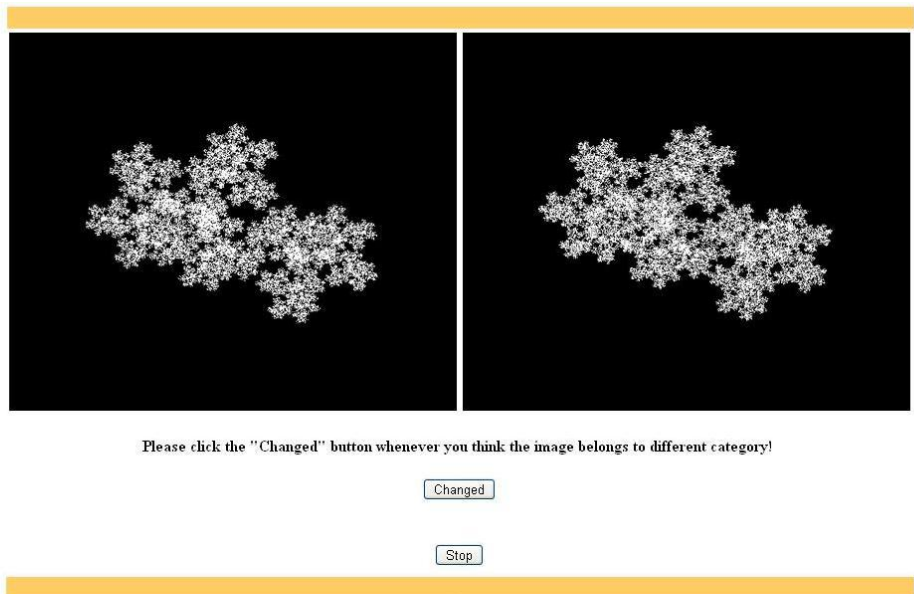


Fig. 3. Screenshot of the web application implemented for the study. This interface shows the sequences of images to the participants, 2 images at a time. New images appear on the right and move to the left.

diagonal, in forward and backward directions. Each session took between 30 and 40 minutes. Any sequence was assigned to only 1 participant, yet almost all images were seen by two participants, except for those images at the intersection of two sequences.

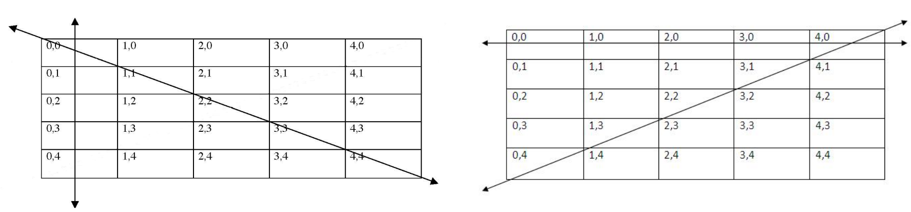


Fig. 4. Method of Limits Directions for Columns and Diagonals

4 Analysis

The results of the computational analysis is shown, in part, in Figure 7. The values for each metric were rescaled so that the minimum was set to 0 and the maximum was set to 1. Four metrics were computed:

- **Fractal Dimension** was determined using the calculator written by Paul Bourke[1].
- **Lacunarity** was computed with a MATLAB[2] routine[3].
- **Mean Squared Error (MSE)** values were computed for images in the same way that participants viewed them: forward and backward. Therefore, for all but the first and last image in a sequence, there were two MSE values, which were then averaged. The first and last image in each sequence were not considered because participants compared them with an image in the next or previous set of sequences which might be totally different.
- **Principle Component Analysis (PCA)** used the eigenface approach of Turk and Pentland [12], laid out in a MATLAB procedure by Dailey[4]. a principle component analysis (PCA) of the stimulus images was performed. The diagonals of the image matrix ($\theta_1 = \theta_2$, and $\theta_2 = 356.4 - \theta_1$) (199 images in total) were used for training. The 40 most important eigenvectors of the covariance matrix for the set of training images formed the basis of a 40-dimensional coordinate system for the images. Values for the image sequences (shown as part of Figure 7) were obtained from the magnitude of the vector defined between the 40-dimensional image coordinates and the origin of that space.

  Once the projection was done in MATLAB, the coordinates were partitioned using the Partition Around Medoid (pam) procedure in R. Figure 5 shows the images clustered into 6 groups, based on their 40-dimensional coordinates. Similar images have the same colour.

Ideally, participants would click at or near the same place when viewing a sequence in forward or backward directions. This did not happen, and clicks were often not even matched between the two viewing directions. Therefore, we treated the clicks as distinguishing between piles of similar images. We had two ratings for each sequence. We derived the pairwise distance for all images in the sequence: if two images were in the same pile, their distance was 0; if two images were in different piles, their distance was 1. Each pair of images could have a distance of 0, 1, or 2 when the two ratings were combined. We used multidimensional scaling to process this data to find where the boundaries between regions of similar images, where clicks were likely to occur, by plotting the distance between neighbouring images. The higher the peak, the more likely

[1] http://local.wasp.uwa.edu.au/∼pbourke/fractals/fracdim/
[2] http://www.mathworks.com/products/matlab/
[3] http://www.mathworks.com/matlabcentral/fileexchange/25261-lacunarity-of-a-binary-image
[4] http://www.cs.ait.ac.th/∼mdailey/matlab/

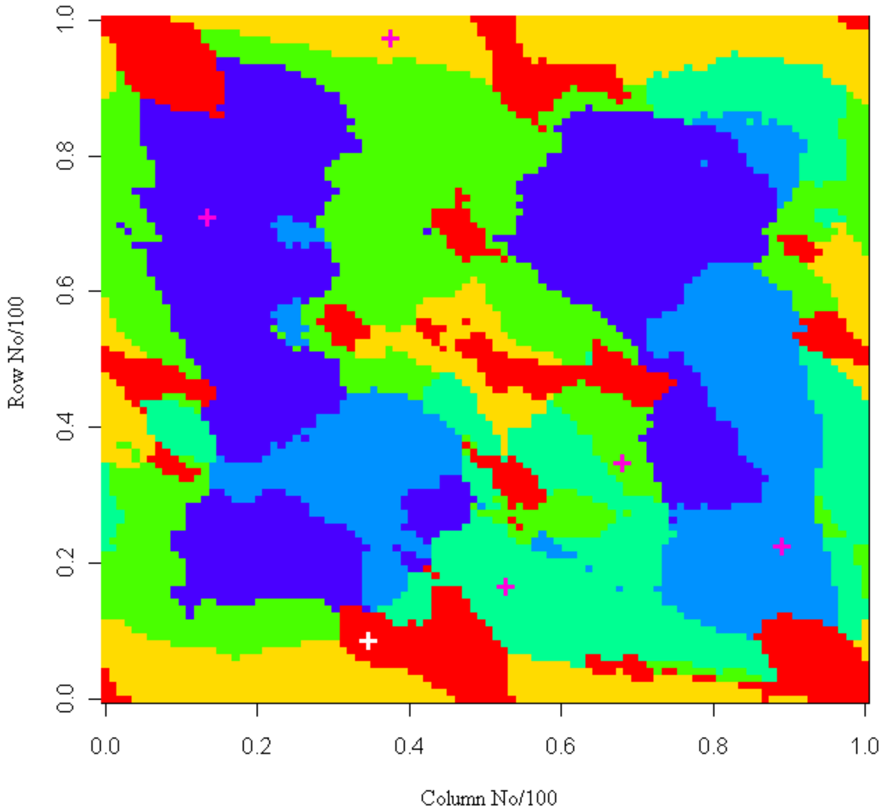


Fig. 5. Clustering of PCA values to 6

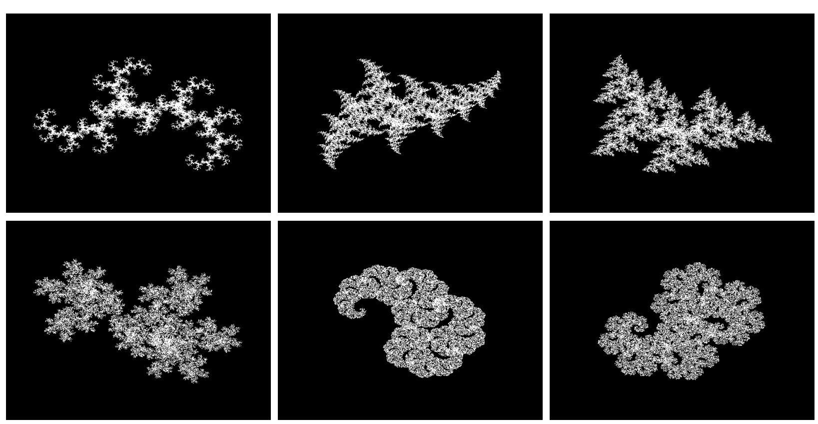


Fig. 6. Fractal images from the centers of each partition illustrated in Figure 5

that a click had occurred there (see Figure 7). The multidimensional scaling was performed with isoMDS routine from R[5].

To enhance the comparison, the values of numerical approaches were computed according to the sequences which the participants saw, as shown in Figure 7. The four vertical lines in Figure 7 show the correspondence between the extrema of the MSE, PCA, fractal dimension, and lacunarity, from left to right with a participant's clicks. Comparing the metrics' values with the participants' clicks shows that most of the clicks corresponds to the extrema in the image comparisons and indicates that participants may be sensitive to them.

To assess the correspondence between participant judgments of similarity (clicks) and computed metrics, we examined a five image window on either side of an extrema value. If there were one or more clicks within that window, one click was identified as corresponding to the extrema. Clicks not in the neighbourhood of any extrema were deemed to be "non-corresponding." Clicks close to the sequence minima (0) and maxima (1) are counted separately. Then, the remaining values are divided into four intervals: $(0, 0.25]$, $(0.25, 0.5]$, $(0.5, 0.75]$, and $(0.75, 1)$.

With the intention of classifying the participants and understanding their behaviour, the participants' corresponding clicks' value along with their total clicks were clustered using the Partition Around Medoid (pam) method in R. A plot of clusters with $k = 2$ is presented in Figure 8. According to the average of their performance, one group seemed most attuned to the fractal dimension and were more accurate in corresponding their clicks to extrema. The other group seemed most attuned to the PCA values, were less accurate with their clicks, and clicked more frequently. This is a very promising start in understanding individual behaviours

5 Conclusion and Future Work

This work has attempted to gain an understanding of how people perceive fractal imagery, and to use that understanding as a first step towards perceptually-based interaction with fractals. The study of linear fractals is apt because the images are straightforward to manipulate and the imagery is beautiful and thought-provoking. This paper has presented new methods for managing and presenting stimuli and for gathering and analyzing similarity ratings of incrementally-varying fractal images. The adaptation of the method of limits was successful and it has opened many questions for further study. The study has indicated that fractal dimension and principal component analysis values may be sufficiently related to participants' judgments of fractal similarity that those computed metrics may be used as approximations of perceived similarity. Some trends seemed to emerge when looking at participant performance on the task, and this seems to be a fruitful area for further investigation. The techniques used in this study may be of help in answering larger questions about the relation between image similarity metrics and perceptual judgements.

[5] http://www.r-project.org/

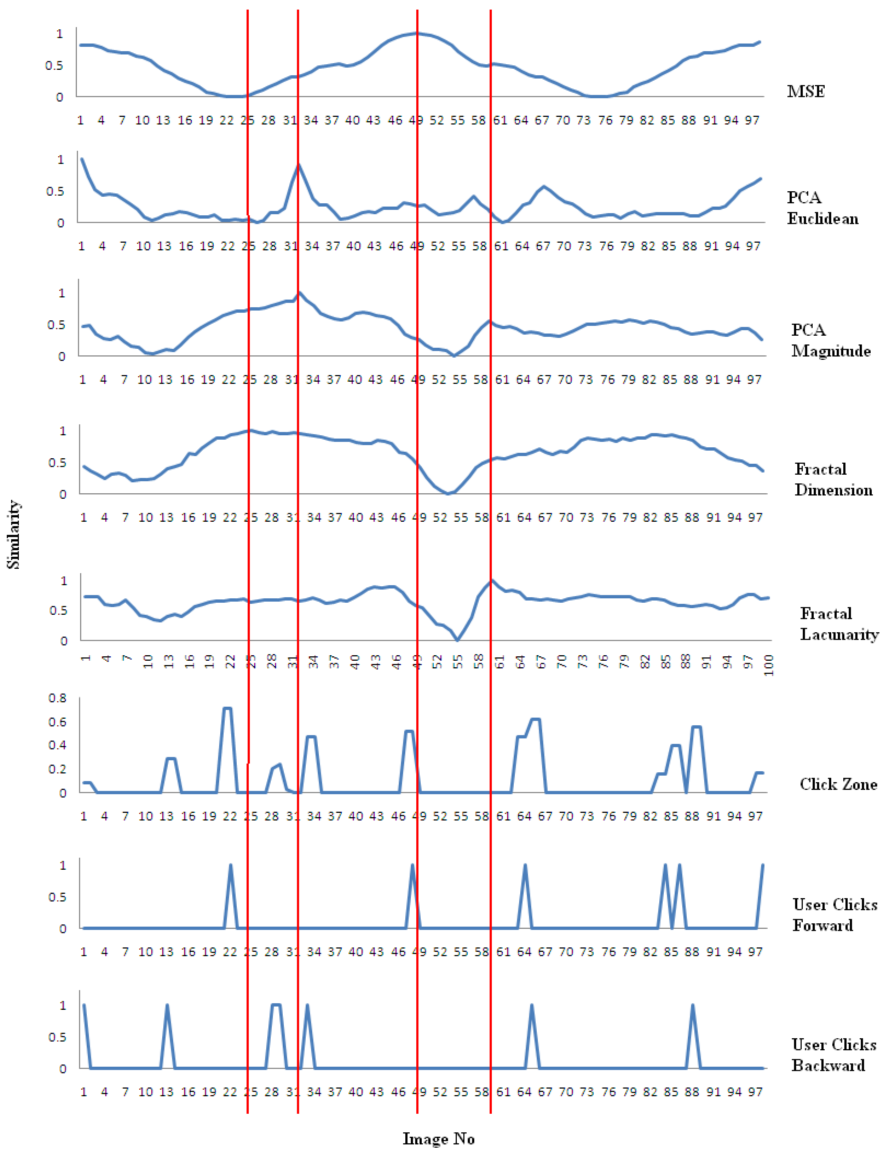


Fig. 7. A comparison of perceptual and computational data, 1st line: MSE values; 2nd line: PCA Euclidean values; 3rd line: PCA Magnitude values; 4th line: fractal dimension values; 5th line: fractal lacunarity values; 6th line: click zone locations, computed from the following 2 lines; 7th line: clicks in forward direction; 8th line: clicks in backward direction. Vertical lines placed at maxima for fractal dimension, PCA, MSE and lacunarity, from left to right.

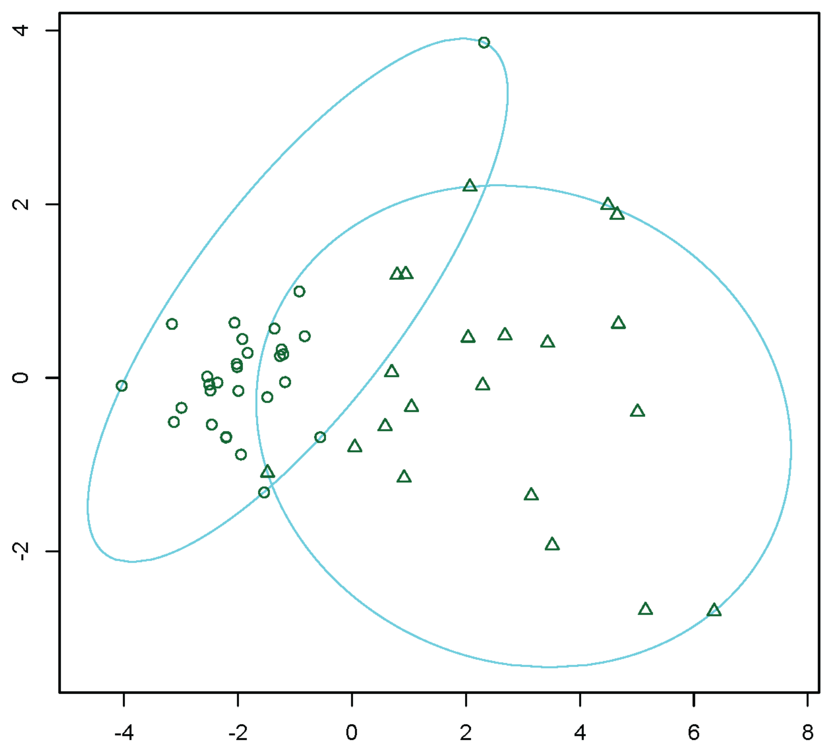


Fig. 8. Clustering the participants into two groups.

The method of data collection could be improved. Many people did complete the task and they were able to identify many extrema indicated by the computational metrics. However, the sessions were too long for some people, and so the quality of the data collected suffered. The design of the experiment emphasized breadth over depth: there was one rating of every image in each direction instead of many ratings of some images. It might be easier to focus on specific areas of the parameter space, which this study has helped to locate.

Furthermore, the operation of the study could be changed. Instead of having the series pass by, perhaps a more effective approach would be to continue to the display the last selected image constant until a new one is clicked.

The method of limits could be applied using different criteria, such as PCA magnitude, fractal dimension, or lacunarity. Further analysis can be done to better understand the relationship between the manipulation of parameters and the features that elicit clicks. There are some indications that users applied different strategies to the task they were given but it would very interesting to study the implications of that possibility for future interface designs. It may be that some of the participants are better informants than others.

References

1. Smith, A.R.: Plants, fractals, and formal languages. ACM SIGGRAPH Computer Graphics **18** (1984) 1–10
2. Whitney, J.: Digital harmony: On the complementarity of music and visual art. McGraw-Hill (1980)
3. Hepting, D.H., Latifi, L., Oriet, C.: In search of a perceptual basis for interacting with parametric images. Proc. Seventh ACM Conference on Creativity and Cognition (2009) 377–378
4. Mandelbrot, B.B.: The Fractal Geometry of Nature. W. H. Freeman (1983)
5. Tel, T., Fulop, A., Vicsek., T.: Determination of fractal dimension for geometrical multifractals. Journal of Physics A **159** (1989) 155–166
6. Plotnick, R., et al.: Lacunarity analysis: a general technique for the analysis of spatial patterns. Physical Review A **44** (1996) 5461–5468
7. Allain, C., Cloitre, M.: Characterizing lacunarity of random and deterministic fractal sets. Physical Review A **44** (1991) 3552–3558
8. Sirovich, L., Kirby, M.: Low dimensional procedure for the characterization of human faces. Journal of the Optical Society of America **4** (1987) 519–524
9. Bonebright, T.L., et al.: Data collection and analysis techniques for evaluating the perceptual qualities of auditory stimuli. ACM Transactions on Applied Perception **2**(4) (2005) 505–516
10. Law, E., von Ahn, L.: Input-agreement: A new mechanism for collecting data using human computation games. ACM Conference on Human Factors in Computing Systems,CHI (2009)
11. Barnsley, M.F., Devaney, R., Mandelbrot, B.B., Peitgen, H.O., Saupe, D., Voss, R.F.: The science of fractal images. (1988)
12. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience **3**(1) (1991) 71–86