# DEVELOPMENT OF A WEB IMAGE BROWSWER THROUGH IMAGE SIGNATURES

Jason M. Kinser
School of Computational Sciences
George Mason University

## ABSTRACT

In recent years web-based information has become more pictorial in nature. However, web search engines are still resigned to matching only text strings. With the onslaught of image data it is now becoming necessary to provide the public with a web image browser. This study will present the use of *image signatures* as a means for representing image data suitable for web image browsing.

**KEYWORDS**: data mining, image signatures, image browsing

## 1.INTRODUCTION

The web has become pictorially oriented as web pages often contain images rather than just text. Unfortunately, our web search engines search generally only for text and not images.

One of the difficulties of browsing for images is that images consume a lot of memory, and therefore, techniques for performing text searches would not be feasible to apply to image searches. The information contained within images first needs to be condensed into a more palatable form. In the mammalian visual cortex this task is performed by converting images into small signatures that are descriptive of the shapes being seen.[1]

In this study, a model inspired by the visual cortex is used to generate image signatures that are descriptive of the shapes within the image. It is possible to find images containing similar shapes by a similarity match of their respective signatures. These signatures are short (fifty elements) and therefore are extremely easy to store and manipulate. A search tree is constructed from signature comparisons and used for an efficient search. The database signature that most closely resembles an unknown signature can be quickly found which indicates which image in the database is most similar (in comparing shapes) to the input image. These tools are the foundation for a web image browsing engine.

## 2.CONSTRUCTION OF IMAGE SIGNATURES

Several models of the visual cortex have been proposed and they contain similar mathematical foundations.[2] Each contains a slab of neurons and each neuron contains at least two coupled oscillators, a non-linear operation, and local connections to neighboring neurons. Instead of choosing a single model, a simple unified model was created [3] that is based on the similarity of the many models and discards the terms that are unique to each. This unified cortical model (UCM) receives an input **S** and assigns one neuron to each input pixel.

Each neuron contains two oscillators, F and $\Theta$, where F is the state of the neuron and $\Theta$ is the dynamic threshold. In this model it is easier to calculate the F's and $\Theta$'s of all the neurons in concert, so these individual oscillators are mathematically grouped in arrays denoted by **F** and Q. The model receives the input and performs iterations over the following equations,

$$\mathbf{F}[n+1] = f\mathbf{F}[n] + \mathbf{S} + W\{\mathbf{Y}\}, \tag{1}$$

$$Y_{ij}[n+1] = \begin{cases} 1 & if \;\; F_{ij}[n+1] > \Theta_{ij}[n] \\ 0 & Otherwise \end{cases}, \tag{2}$$

$$\Theta[n+1] = g\Theta[n] + h\mathbf{Y}[n+1], \tag{3}$$

Here the output Y for any neuron is either a 1 or 0 depending on the comparison of the state and the threshold. Scalar $g$ must be less than $f$ and both are less than one. Scalar $h$ is quite large.

From each iteration the UCM provides an array **Y** which are the output values. Integration of neural activity and the integration of an edge-enhanced version of **Y** are used to compute the signature,

$$G[n] = \sum_{i,j} Y_{ij}[n], \tag{4}$$

$$G[n+1] = \sum_{i,j} \left( Z\{\mathbf{X}\} \right)_{ij} . \quad (5)$$

where Z{} is an edge enhancing function such as the Laplacian.

The first half of the signature follows the model proposed by Johnson [4] and the second half was determined to be useful through modeling.[5]

In previous models the weighted connections between neurons were usually positive and inversely proportional to the distance between neurons. It was determined earlier [6] that these types of connections produced autowave communications. Autowaves [7] are propagations that do not reflect or refract. However, they emanate from a source and continue onward until they collide with a boundary or surface. In the UCM model this type of communication led to *interference*. The presence of one object could drastically alter the neural activity caused by another. In short, the signature produced by one object was not repeatable if other objects existed in the image.

Circumvention of the interference problem was through the realization that autowaves and *curvature flow* [8] models provide similar propagation except for a scale factor. So, W{ } is a function that is sensitive to the shape of the neural pulses in the neighboring vicinity. The output **Y** generally contains solid shapes, which are segments inherent in the original image. Neural activity around the border of these shapes encourages neurons to pulse if they are located towards the center of curvature of the segment. Left alone each shape would eventually morph to a circle and then collapse to a point.[3,9]

## 3.DATABASE

In order to examine the ability of signatures as an effective recognition representation a database was constructed. One thousand images from random web pages were used. This provided a database with several different types of images of differing qualities. However, since some web pages had several images dedicated to one topic the database contained sets of similar images. There were even a few exact duplicates and several duplicates of differing scale. The only qualifications of the images was that they had to be of sufficient size (more than 100 pixels in both dimensions) and sufficient variance in intensity (to prevent banners from being used).

The database itself, consisted of the signatures, the original URL, and the data retrieved. It was not necessary to keep the original images. Each image thus required less than two hundred bytes except for cases of very lengthy URLs.

## 4.CLASSIFICATION

Comparison of the signatures was accomplished through a normalized subtraction. Thus, the scalar representing the similarity of two signatures $G_q$ and $G_p$ was computed by,

$$a = 1.0 - \sum_n \left| \left( \left\| G_p[n] \right\| - \left\| G_q[n] \right\| \right) \right| . \quad (6)$$

The signatures were normalized to eliminate the effects of scale.

### 4.1 BEST MATCHES

The signatures of each in the database were compared to find the best matches. Since there were 1000 images in the database there were 499,000 different possible pairings (excluding self-pairings). The top scores were listed and the images were manually compared.

In the data base there eight duplicate images (each pairing scored a perfect 1.0 by equation 6). There were 11 pairings that scored above 0.9447 and each pairing was a single image at different scales. Other scores for pairings of images with different scales were 0.9408 and 0.9126. One pairing was two versions of the same images with a scale factor of 1.43 in the vertical and 1.29 in the horizontal. It scored 0.9016.

The rest of the top scores consisted of pairs that could be placed in three groups. Pairs of images that contain very similar objects but were not duplicate images scored 0.9380, 0.9175, 0.9163, 0.9117, 0.9099, 0.9085, and 0.9081. Pairs that were somewhat similar provided scores of 0.9223, 0.9123, 0.9117, 0.9098, 0.9083, 0.9077, 0.9065, and 0.9062. Finally, the top scoring pairs that seemed to have very little in common with the images were 0.9433, 0.9204, 0.9120, and 0.9088.

It was possible to mostly separate the perfect matches, the scaled pairings, and the similar images from the rest. This, of course, was not an exhaustive study since it was time consuming.

### 4.2 ONTOLOGY

In order to increase the ability to compare signatures and analyze the results an ontological tree was constructed. The best pairs according to equation 6 created a branch in the tree. The top of each branch was replaced with the average of the two signatures and an entire tree could be constructed to sort the signatures according to similarity.

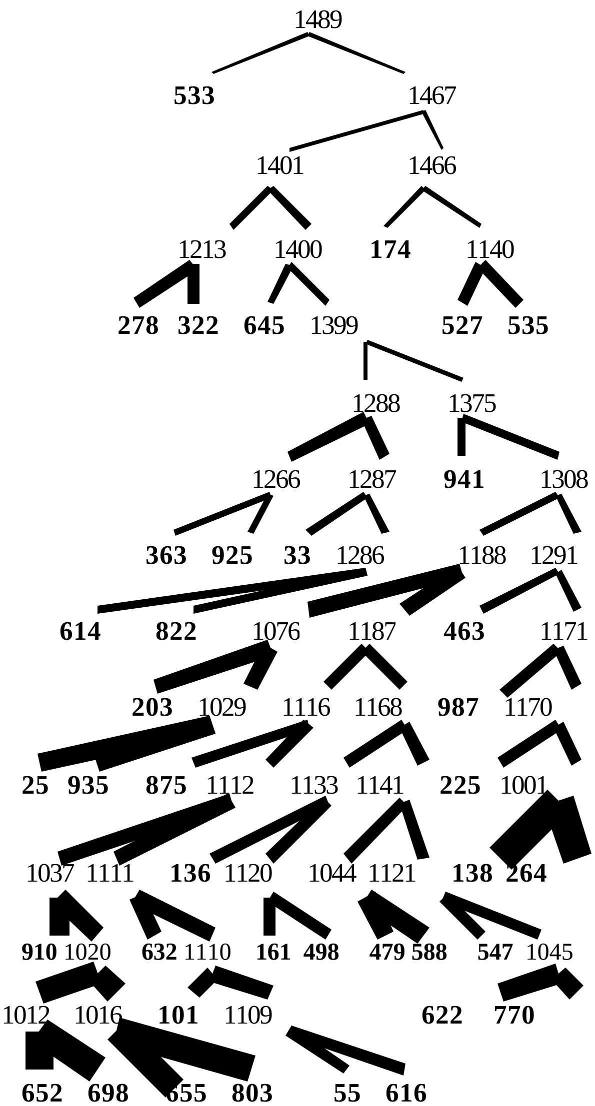


*Figure 1. Partial tree structure.*

A partial segment of the tree is shown in figure 1. The entries with values less than 1000 correspond to the original image number. The nodes with values greater than 1000 are the new signatures created from averaging. Thus, #1012 is the average of original images #652 and #698. The thickness of the connecting lines is proportional to the score from equation 6.

The image in figure 2 displays a portion of the tree shown in figure 1 with the images. It is rearranged a bit for presentation. The Olympic flames are image #652 and #698. The flowers are #655 and #803. The photo of the people in the water is #910, the collage is #632, and interior of a building #101, the soldiers #55, and the diner #616. The thin line links are scores that are below 0.900 and therefore well within the range of dissimilar image scores. However, it was the best match for that image within this database.

## 4.3 SEARCH

Searching the database starts at a top of the tree and works down through the best scoring nodes. However, since many of the links in the tree are through small scores this is difficult to accomplish in a single tree. So, the single large tree was dissected into smaller trees in which all of the links were above a threshold $\gamma$. As the search moved down each tree a node survived only if the score was greater than $\phi * p$ where $\phi$ is a threshold scalar and p was the previous score of that node. Figures 3 and 4 show the number of trees and the depth of the largest tree as a function of the threshold $\gamma$. Thus for $\gamma$=0.8 the original tree structure was separated into 435 individual trees, and the largest depth (the number of nodes from the top of the tree to the last node) was 18. More efficiency is gained in the search by having the fewest number of trees since each tree requires its own search.

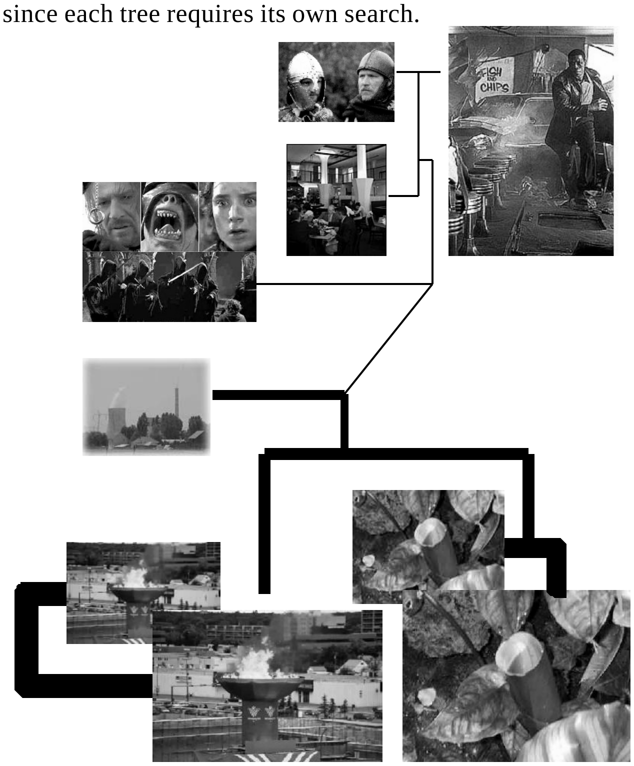

*Figure 2. Partial tree with images.*

To test this tree each of the original images was used as a probe. A match was successfully found if the probe image exactly matched the image provided by the search. Figures 5 and 6 show the number of recall errors versus the thresholds. There are 1000 tests and so 10 errors is 1% of the total recall. As $\gamma$ increases the number of errors falls. This indicates that there is a $\gamma$ at which the connections between nodes exceeds the comparison capability of equation 6. The threshold $\phi$ determines which nodes should survive during a search. Of course, fewer computations are required if nodes can be readily pruned. The trade-off is that the node that contains the correct answer should be not pruned in the earlier levels of the search. The chart in figure 6 displays at which erroneous pruning begins.

For this database the optimal search was $\gamma$=0.75 and $\phi$=0.95. These parameters are dependent upon the content of the database.

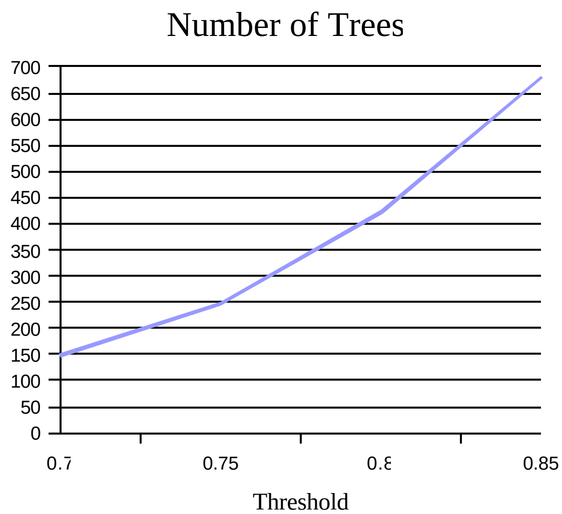


*Figure 3. Number of trees vs. g.*

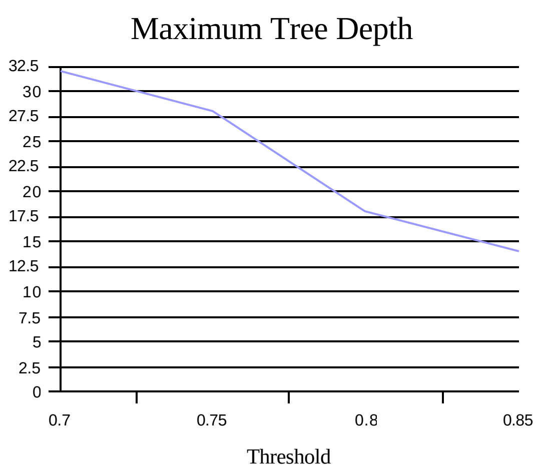


*Figure 4. Size of the largest tree vs. g.*

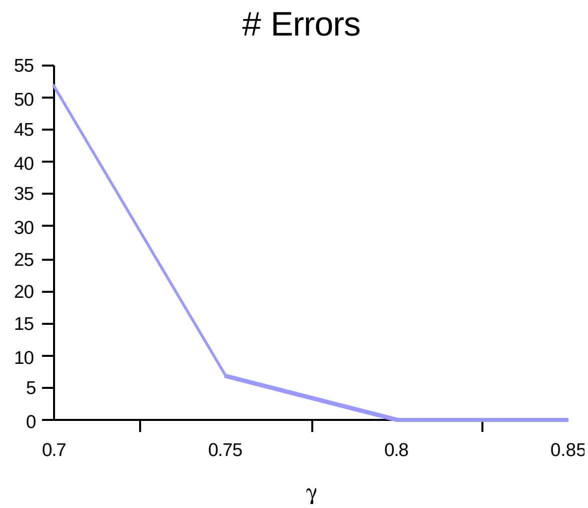


*Figure 5. Number of errors vs. $\gamma$.*

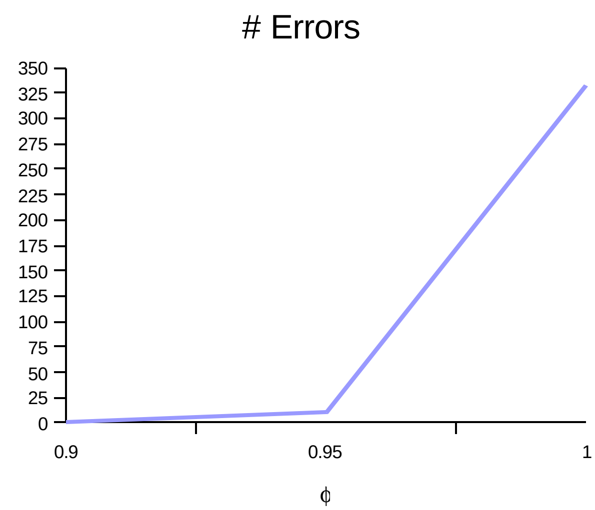


*Figure 6. Number of errors vs. $\phi$.*

## 4.4. WEB IMAGE BROWSING

The search of a large database is not much different than searching the web. As crawlers come across the images the signatures are computed and stored into a tree organization. As the number of images grows it is expected that eventually these minor trees will be connected to make larger trees. This will increase the efficiency of the search since the number of comparisons is strongly related to the number of trees.

## 5.CONCLUSIONS

In the pursuit of a web image browser it is necessary to condense the pertinent information into small data streams. Image signatures provide a condensation of data that is about three orders of magnitude. These signatures can be used for efficient database searches designed to find the database image that is most similar to a probe. Early studies indicate that efficient tree structures consisting of these signatures can be built and searched.


## 6. ACKNOWLEDGEMENTS

This project was funded, in part, by a University Research Initiative grant from the National Imagery and Mapping Agency. The author gratefully thanks NIMA for their support of this project.


## REFERENCES


[1] J. W. McClurkin, J. A. Zarbock, L. M. Optican, “Temporal codes for colors, patterns and memories”, *Cerebral Cortex* **10**, 1994, 443-467.

[2] J. M. Kinser, “Hardware: Basic Requirements for Implementation”, Proc. Of SPIE, **3728**, Stockholm, June 1998, 222-229

[3] J. M. Kinser, C. Nguyen, “Image object signatures from centripetal autowaves”, *Pattern Recognition Letters*, **21**(3), 221-225, (2000).

[4 ]J. L. Johnson, “The signatures of images”, *IEEE Intl. Conf. Neural Networks*, Orlando, 1994.

[5] J. M. Kinser, “Image Signatures: Ontology and Classicification”, submitted Fourth IASTED International Conference on Computer Graphics and Imaging (CGIM 2001) (2001).

[6] T. Lindblad and J. Kinser, *Image Processing using Pulsed Coupled Neural Networks*, Springer-Verlag, London, 1998.

[7] O. A. Mornev, “Elements of the ‘Optics’ of Autowaves” in Self-Organization Autowaves and Structures far from Equilibrium, V. I. Krirsky ed., Springer-Verlag, pp111-118 (1984).

[8] M. A. Grayson, “The Heat Equation Shrinks Embedded Plane Curves to Round Points”, *J. Differential Geometry* **26**, 1987, 285-314.

[9] J. M. Kinser, “Hardware: Basic Requirements for Implementation”, Proc. Of SPIE, **3728**, Stockholm, June 1998, 222-229.