

IMAGE SIGNATURES: ONTOLOGY AND CLASSIFICATION

Jason M. Kinser
School of Computational Sciences
George Mason University
Manassas, VA

ABSTRACT

As our ability to generate a plethora of image data increases it becomes important to develop tools that are capable of searching image databases for similarities. This requires the extraction of content or shape information rather than relying on algorithms that perform pixel matching. In order to accomplish this task this study converts images into signatures which are short vectors that are descriptive of the shapes contained within the image. Similarity measures of these vectors indicate the similarity of shapes contained within the two images. In this study an ontology of more than 1000 image signatures was constructed and the search results are presented.

KEYWORDS: image signatures, pulse image processing, image ontology

1. INTRODUCTION

Image signatures are short vectors (less than 100 elements) that are descriptive of the shapes contained within an image. Since they are perhaps three to five orders of magnitude smaller than the original image the signatures are far easier to organize, store, and search. The most important quality of the signatures is that they are descriptions of the shapes within the image and thus similarity matching of the signatures concurrently finds images containing similar shapes. Since image content is usually represented by shapes the image signatures are powerful representations of the image.

In this study, the theory of image signatures will be presented. A simple metric for measuring signature similarity will be employed to find the most similar signatures. Then signatures from 1000 images extracted from random web sites will be used to build an ontology that groups similar signatures. Finally, this ontology will be used as a efficient search tool in

order to find the best match for an unknown amongst the entire database.

2. IMAGE SIGNATURE THEORY

Image signatures are derived from models of the mammalian visual cortex. There is evidence that mammals create image signatures [1] although the mammalian mechanisms are not yet thoroughly understood. The digital model does depart from biological models to improve simulation performance and increase computational efficiency.

2.1. BACKGROUND

Roughly a decade ago researchers were presenting models of the mammalian visual cortex. [2,3,4] Digital simulations revealed algorithms that were capable of extracting segments that were inherent in the original image. [5,6,7] Each neuron had at least two coupled oscillators, a non-linear function, and positive local connections.

2.2. THE UCM

For computer simulation a model was developed that contained these similar qualities and discarded the terms unique to individual models. [8] Each neuron consists of two coupled oscillators, a non-linear operation, and local connections. The state of the neuron is F , the threshold is Θ , and the output is Y . A single iteration is described by,

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

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

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

where \mathbf{F} is the array representing the state variables from all neurons. Likewise, \mathbf{Y} and Θ are arrays representing the output and threshold for all

neurons. The variables f and g are less than 1 and $g < f$ in order for the threshold to decay faster than the state, otherwise, the neurons would never pulse. The variable h is quite large and gives the threshold a big boost when the neuron fires. The function $W\{ \}$ describes the inter-neuron communications. In order to eliminate *interference* [9] caused by connections in earlier proposed models, this system propagates the neural information as centripetal *autowaves* [9] which are propagating waves that do not reflect or refract.[10] Instead of traveling away from the origin, centripetal autowaves propagate towards to local center of curvatures of the wavefront. This is similar to curvature flow theory. [11,12]

A simple curve, C , evolves such that the boundary moves towards the local center of curvature $\frac{\partial C}{\partial t} = \kappa \cdot \hat{n}$, where κ is the curvature vector for any point on the curve and \mathbf{n} is normal. In two-dimensional space all shapes become a circle and then collapse to a point.

For the creation of image signatures it is important to capture qualities of shape. As the autowaves and propagate they relay information about the shape of the pulsing region, and the boundaries from two different shapes will not interfere.

The computation of $W\{*\}$ creates centripetal autowave communications. Computations for curvature can be cumbersome for large images. So, an image-friendly approach following the procedures of [12] is adopted.

The function $W\{A\}$ is computed by

$$W\{\mathbf{A}\} = \mathbf{A}' = \left[\left[F_{2,A}\{M\{\mathbf{A}'\}\} + F_{1,A}\{\mathbf{A}'\} \right] < 0.5 \right], \quad (4)$$

where

$$\mathbf{A}' = \mathbf{A} + \left[F_{1,A}\{M\{\mathbf{A}\}\} > 0.5 \right]. \quad (5)$$

The function $M\{\mathbf{A}\}$ is a smoothing function. The function $F_{1,A}\{\mathbf{X}\}$ is a masking function that allows only the pixels originally OFF in A to survive as in,

$$\left[F_{1,A}\{\mathbf{X}\} \right]_{ij} = \begin{cases} X_{ij} & \text{if } A_{ij} = 0 \\ 0 & \text{Otherwise} \end{cases}, \quad (6)$$

and likewise $F_{2,A}\{\mathbf{X}\}$ is the opposing function,

$$\left[F_{2,A}\{\mathbf{X}\} \right]_{ij} = \begin{cases} X_{ij} & \text{if } A_{ij} = 1 \\ 0 & \text{Otherwise} \end{cases}. \quad (7)$$

The operators $>$ and $<$ are threshold operators as in,

$$\left[\mathbf{X} > d \right]_{ij} = \begin{cases} 1 & \text{if } x_{ij} \geq d \\ 0 & \text{Otherwise} \end{cases}, \quad (8)$$

and

$$\left[\mathbf{X} > d \right]_{ij} = \begin{cases} 1 & \text{if } x_{ij} \leq d \\ 0 & \text{Otherwise} \end{cases}. \quad (9)$$

This system works by basically noting that a smoothed version of the original segment produces larger values in the off-pulse region and lower values in the on-pulse region in the same areas that the front is to propagate. The nonlinear function isolates these desirable pixels and adjusts the communication wavefront accordingly.

2.3. Image Signatures

The first attempt to generate image signatures [13] was to simply integrate neural activity for each iteration (equation 10). For the case of very simple objects it was shown that the signatures were invariant to shift and rotation. They were also invariant to a limited range of scale, skew, and illumination alterations. However, this method did not work for non-trivial images due to interference.

Even with the new model that removed interference the signatures were strongly linked to object size and weakly to shape. This was insufficient for recognition by signatures for larger databases. Thus, a second part of the signature was created. The first part was the original integration and the second was an integration of an edge-enhanced version of the pulsing patterns. Thus, the signature was computed by,

$$G[n] = \sum_{i,j} Y_{ij}[n], \quad (10)$$

$$G[n+1] = \sum_{i,j} (Z\{\mathbf{X}\})_{ij}. \quad (11)$$

An example is shown in figures 1 and 2.



Figure 1. An image.

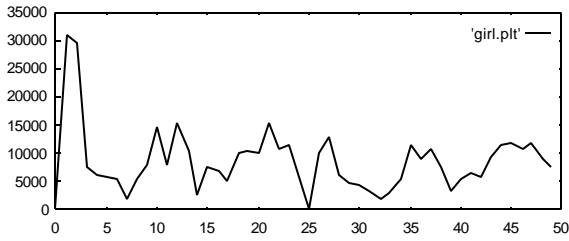


Figure 2. The signature of the image in figure 1. The first 25 elements come from equation 10 and the last 25 elements from equation 11.

3. COMPARING SIGNATURES

The image signatures describe shape information and thus similar signatures should indicate similarity of shapes contained within images. This, however, is a subjective examination since image similarity is in the eye of the viewer and not rigorously defined.

3.1. Metric

Two signatures are similar if the difference between their normalized signatures is small. The comparison of two signatures G_p and G_q is computed by,

$$a = 1.0 - \sum_n \left(\left| \|G_p[n]\| - \|G_q[n]\| \right| \right). \quad (12)$$

The maximum score of $a=1$ indicates that there was a perfect match.

3.2. The Database

The database consisted of 1000 images extracted from random web sites. Each image was of sufficient size (more than 100 pixels in either dimension) and had sufficient variation in pixel intensities (to reject banner and cartoon images). Images that were large were reduced and all images were converted to gray scale.

3.3. Similarities

In order to view the ability to distinguish signatures one signature was compared to the entire database. The plot in figure 3 depicts the histogram of image similarity. The original image and the closest matches are shown in figure 4. The windows in the original image find similarities to the other images. Recall that these are the best matches, not necessarily good matches.

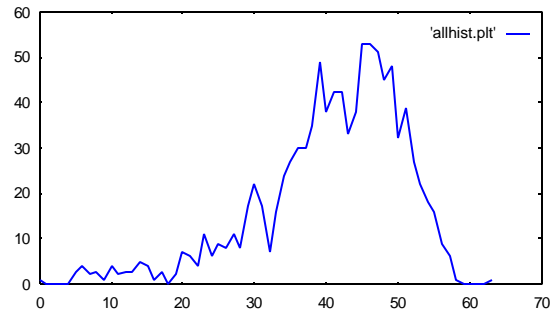


Figure 3. Histogram of matches of all entries in image #966 (the image with the largest total score). The x axis is the number of bins.



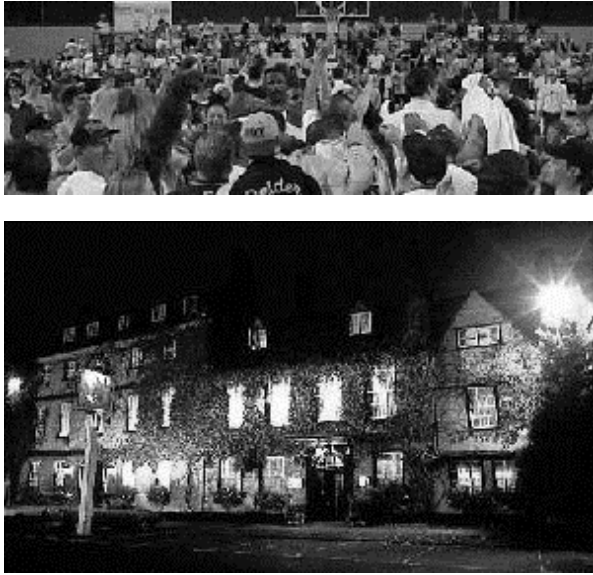


Figure 4. Closest matches. The scores from the first images to all of the others is 0.87.

4. ONTOLOGY

An ontology is a structured organization of the data. In the search for an image similar to an unknown the ontology structure should provide an efficient search space. Thus, the number of searches should be less than the number of entries in the database.

4.1. Tree Construction

A well known method of organizing vectors into a binary tree is the UPGMA (unweighted pair group method with arithmetic mean).[14] In the original UPGMA algorithm a single tree would eventually be built that connected all of the vectors independent of strength of those connections. Here the UPGMA is used to construct an image signature ontology. This is an undesired quality for the employment of the tree as a search engine since some of the connection can be very weak. Thus, a threshold, γ , was established. Connections were not made if $a < \gamma$. The end result was that there were several smaller trees rather than one large tree. To begin the search the top node of each tree had to be compared to the unknown signature.

The number of trees with respect to the threshold value is shown in figure 5. In figure 6 the number of levels in the largest tree are plotted with respect to the threshold value. Of course, the desire is to have a fewer number of trees to reduce the number of initial comparisons that must be made.

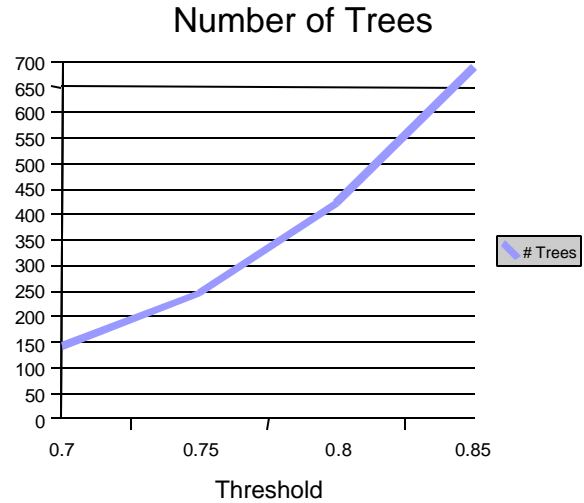


Figure 5. The number of trees vs. γ .

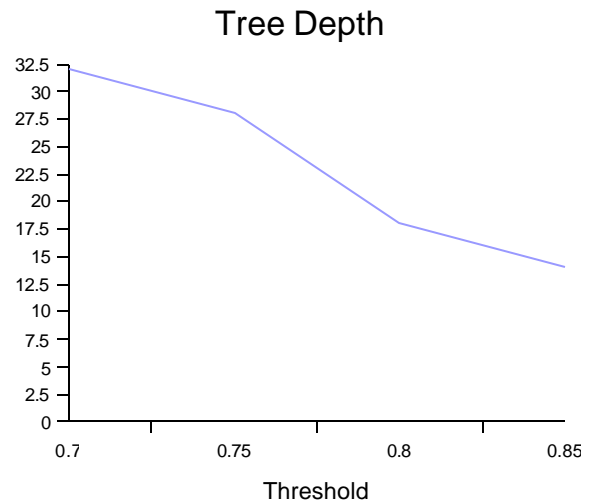


Figure 6. Depth of the largest tree vs. γ .

4.2. Recall

The recall is performed by comparing the signature of the unknown to the top nodes in all the trees. Subsequent branches of each node survived if their score was greater than ϕ multiplied by the node's score. The graph in figure 7 plots the number of searches needed by the algorithm versus γ and the graph in figure 8 plots the number of searches needed versus f for a constant γ of 0.8. The UPGMA method creates 999 new nodes as each and thus a single tree would contain 1999 nodes. It is possible, with improper settings to actually perform more comparisons than number of images in the database.

The goal is to set ϕ and γ to minimize the number of nodes of actually considered during a search without infusing recall errors.

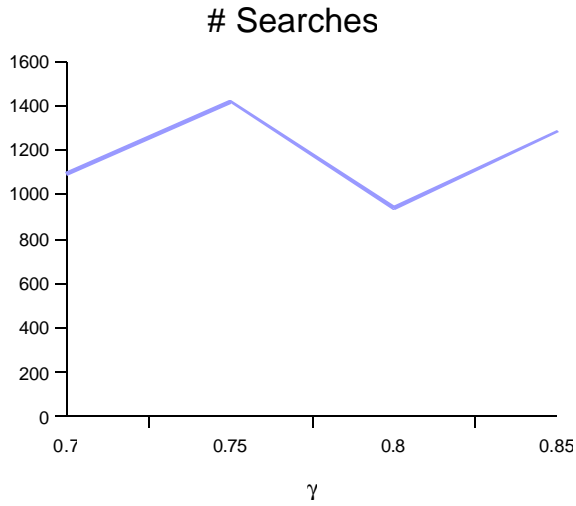


Figure 7. Number of searches vs. γ .

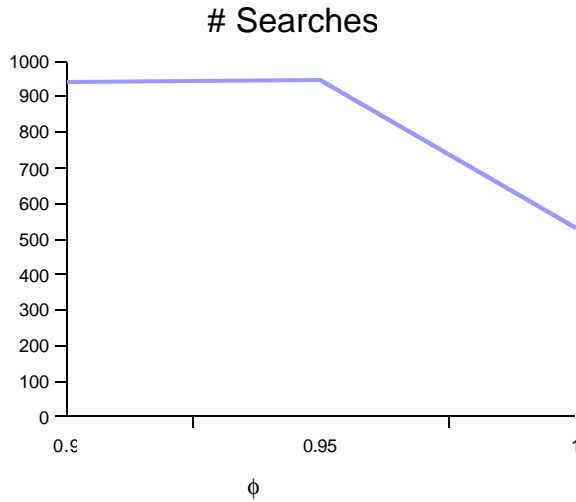


Figure 8. Number of searches vs. ϕ .

Finally, the performance of the system needs to be addressed. In the test each of the signatures of the database were used as an unknown. Thus, each unknown should have an exact match in the database. In some cases, the input signature was identical to two signatures in the database and so the identification of either of these database signatures was considered a success. The graph in figure 9 plots the recall errors versus γ and the graph in figure 10 plots the recall errors versus ϕ for a constant γ of 0.8.

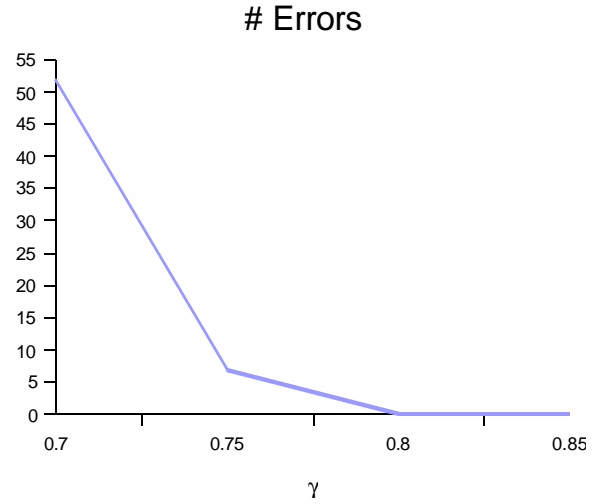


Figure 9. Number of errors vs. γ .

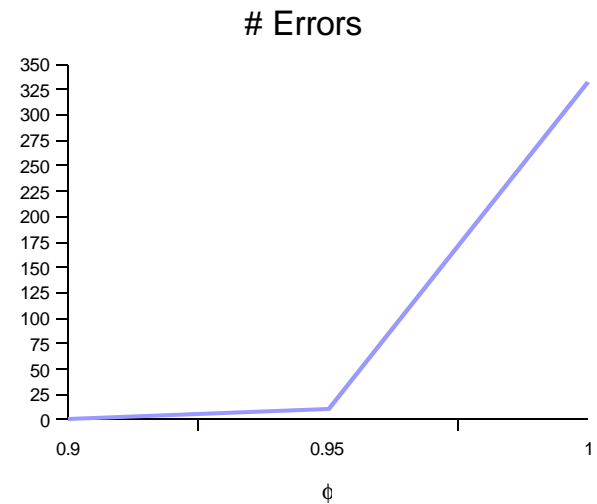


Figure 10. Number of errors vs. ϕ .

5. CONCLUSIONS

Image signatures can be used for highly efficient representations of the collection of shapes within an image. Research is continuing on finding methods that improve search efficiency. The advantages so far presented by this method are twofold. The first is that this method image content rather than performing pixel matching or histograms. The information used is much closer to that used by humans. The second is that the amount of information used is reduced to fifty integers for each image. This reduction could be as much as five orders of magnitude for large images.

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, *Cerebral Cortex* **10**, 1994, 443-467.
- [2] R. Eckhorn, H. J. Reitboeck, M. Arndt, P. Dicke, *Neural Computation* **2**, 1990, 293-307.
- [3] I. A. Rybak, N. A. Shevtsova, V. A. Sandler, *Neurocomputing* **4**, 1992, 93-102.
- [4] O. Parodi, P. Combe, J-C. Ducom, *Biol. Cybern.* **74**, 497-509 (1996).
- [5] J. L. Johnson, H. S. Raganath, H. J. Caulfield, *IEEE Conf on Neural Networks*, Orlando FL, 1994.
- [6] Z. Li, *Network: Comput. Neural Syst.* **10** 187-212, (1999)
- [7] T. Lindblad and J. Kinser, *Image Processing using Pulsed Coupled Neural Networks*, Springer-Verlag, London, 1998.
- [8] J. M. Kinser, *Proc. Of SPIE*, **3728**, Stockholm, June 1998, 222-229.
- [9] J. M. Kinser, C. Nguyen, *Pattern Recognition Letters*, **21**(3), 221-225, (2000).
- [10] 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).
- [11] M. A. Grayson, *J. Differential Geometry* **26**, 1987, 285-314.
- [12] R. Malladi and J. A. Sethian, *Proc. of Conf. On Visualization and Mathematics*, June 1995, Springer-Verlag, 329-343.
- [13] J. L. Johnson, *IEEE Intl. Conf. Neural Networks*, Orlando, 1994.
- [14] C. D. Michener, R. R. Sokal, *Evolution*, **11** (1957) 130-162.