

Foveation by a Pulse-Coupled Neural Network

Jason M. Kinser
Center for Applied Optical Sciences
Dept. of Physics, PO Box 1268
Alabama A & M Univ.
Normal, AL 35762
jaker@mycroft.caos.aamu.edu

Keywords

Foveation, Pulse-Coupled Neural Network, Visual Cortex, Target Recognition

ABSTRACT

Humans do not stare at an image, they foveate. Their eyes move about points of interest within the image collecting clues as to the content of the image. Object shape is one of the driving forces of foveation. These foveation points are generally corners and, to a lesser extent, the edges. The Pulse-Coupled Neural Network (PCNN) has the inherent ability to segment an image. The corners and edges of the PCNN segments are similar to the foveation points. Thus, it is a natural extension of PCNN technology to use it as a foveation engine. This paper will present theory and examples of foveation through the use of a PCNN, and it will also demonstrate that it can be quite useful in image recognition.

1. INTRODUCTION

The human eye does not stare at an image. It, in fact, moves to different locations within the image to gather clues as to the content within that image. A typical foveation pattern [Yarbus,65] is shown in Fig. 1. Many

of the foveation points are on corners and edges of the image. More foveation points indicate an area of greater interest.

The Pulse-Coupled Neural Network (PCNN) [Johnson,94] has the inherent ability to segment images. The corners and edges of these segments are similar to many of the foveation points that humans naturally use. Couple this with the fact that the PCNN is based on a model of the visual cortex of a small mammal [Eckhorn,90] and it becomes a good candidate to simulate foveation.

This paper will then present a foveation algorithm based strongly on the PCNN and results from that algorithm. It will also present a simple image recognition system based upon foveation points derived by this method. Since the PCNN is a relatively new technology it will also be quickly reviewed.

2. The Pulse-Coupled Neural Network

The typical neuron of the PCNN is shown in Fig. 2. It is more complicated than the traditional McCulloch-Pitts neuron. The PCNN neuron has a feeding and linking input. These are then combined in a second order fashion and then compared to a dynamic threshold. The equations for a single iteration of the PCNN are,

$$F_{ij}[n] = e^{-\tau_F} F_{ij}[n-1] + S_{ij} + V_F \sum_{kl} M_{ijkl} Y_{kl}[n-1], \quad (1)$$

$$L_{ij}[n] = e^{\Delta t \Delta_i} L_{ij}[n-1] + V_L \sum_{kl} W_{ijkl} Y_{kl}[n-1], \quad (2)$$

$$U_{ij}[n] = F_{ij}[n] \{1 + \Delta L_{ij}[n]\}, \quad (3)$$

$$Y_{ij}[n] = \begin{cases} 1 & \text{if } U_{ij}[n] > \Delta_{ij}[n] \\ 0 & \text{Otherwise} \end{cases}, \quad (4)$$

$$\Delta_{ij}[n] = e^{\Delta t \Delta_i} \Delta_{ij}[n-1] + V_{\Delta} Y_{ij}[n], \quad (5)$$

where F is the feed, L is the link, U is the internal activity, Y is the pulse output and Δ is the dynamic threshold. The local connections \mathbf{M} and \mathbf{W} are fixed (usually Gaussian). This system requires no training.

Through these local connections the activation of a neuron adds to the internal activity of the neighboring neurons. Groups of neurons receiving similar stimulus that are spatially close to each other tend to synchronize pulses. This is the foundation of the inherent segmentation ability of the system.

3. The Foveation Algorithm

The foveation algorithm relies heavily on the segmentation ability of the PCNN. The segments produced by the PCNN are then filtered to extract

the foveation points. The logical flow of the system is shown in Fig. 3. Basically, the PCNN produces a series of binary pulse images. These images contain a few segments from the original image. Each frame contains a different set of segments. These segments are sent through a low pass filter that enhances the appropriate areas. Examples of a pulse image and its filtered image are shown in Figs. 4. The effect of filtering is as desired. The corners and some of the edges of the larger areas become significantly brighter than the interior, medium sized areas become a smooth area with a single peak, and the smaller areas are reduced to a level of insignificance. Finding the foveation areas now becomes a matter of peak detection which is quite simple since each of the peak areas is quite smooth.

Each frame was searched for peaks that were within 90% of the initial peak value. These were then reserved for further processing which will be discussed later.

The first example is that of the handwritten letters. Foveation points for these letters are corners and junctions. The letters in Figs. 5 shows the original letters and the foveation points. These points are ranked to indicate the order in which the foveation points were found.

The second example is that of the face from Fig. 1. Unfortunately, the PCNN does not work well with binary inputs and the original face image is binary. So this image was smoothed to give it texture which destroyed some of the smaller features such as the eye. The foveation points selected by the PCNN model are shown in Fig. 6. It should also be noted that the PCNN algorithm is not simulated all of the sources for foveation. Humans are “hard-wired” to recognize humans and foveation on an image of a face is

driven by more than just the shapes contained within the image. However, there encouraging similarities between the some of the human foveation points and the points computed by the PCNN model.

The final example is to compute the foveation points for a fairly complicated image with many objects, noise, and background. The iamge and the its foveation points are shown in Fig. 7. Many of the foveation points were along lines and edges of within the image. However, lesser important details such as the car grill and the features on the boy's shorts did not produce foveation points. It should also be noted that low contrast larger features (such as the edges between the bottom of the car bumper and the ground) did produce foveation points. Thus, it can be seen that these foveation points are quite different than those that would have been produced by edge filtering the original image.

Unfortunately, it is not possible to say which foveation points are correct - or least which ones mimic humans. However, the PCNN model does produce foveation points that are desired which are regions of corners and edges.

4. Target Recognition by a PCNN Based Foveation Model

This section presents preliminary results of a handwritten character recognition test that is strongly based upon a foveating PCNN. In general, the foveation points were selected and filtered. Thus, the algoirthm is capable of attending to important features and recognizing them. The identification of a

target based upon fuzzy scoring of identified features has already been published. [Srinivasan,96]

Typical handwritten characters are shown in Fig. 7. In this database there were 7 samples from a single person and 1 sample of 3 letters (A, M and U) from 41 individuals. The foveation of these points are typically shown in Fig. 5. The logic flow of the recognition system is shown in Fig. 9.

Once the foveation points are produced a new images are created by a barrel transformation centered on each foveation point. An example of the letter 'A' and barrel transformations centered on the foveation points are shown in Fig. 10. This distortion places more emphasis on the information closer to the foveation point. Recognition of these images constitutes the recognition of a feature of the image. It should take a few features to recognize an object.

Recognition of the feature images is performed through a Fractional Power Filter (FPF) [Brasher,94]. This filter is a composite filter that has the ability to manipulate the trade-off between generalization and discrimination that is inherent in first order filters. In order to demonstrate the recognition of a feature by this method an FPF was trained on 13 images of which 5 were target features and 8 were non-target features. For this example the target is top of the 'A' (see Fig. 10b).

The results of the test are presented as three categories. The first measures how well the filter recognized the targets, the second is how well the system rejected non-targets, and the third considers the case of a non-target that is similar to the target (such as the 'M' has two features similar to the

top of the 'A'). The maximum correlation signature about the area of the foveation point were recorded. The FPF was trained to return have a correlation peak of 1 for targets and a peak of 0 for non-targets. The results for (non-training) targets and dissimilar non-targets are shown in Table 1. Similar non-targets produced significant correlation signatures as expected. Certainly, an object can not be recognized by a single feature.

The results demonstrate an exceedingly good separation between targets and non-targets. There were a few targets that were not recognized well. Such target come from Fig. 8i, Fig. 8j and Fig. 8k. Fig. 8i is understandable since the object is poorly represented. Fig. 8j performed poorly since the top of the 'A' is extremely narrow, and Fig. 8k has an extremely rounded feature and the top. These last two features were not represented in the training features. All othe types of 'A's produced a correlation signature above 0.8 which is clearly distinctive from the non-targets.

A few false negatives are not destructive to the recognition of the object. Following the example of [Srinivasan,96] a collection of features recognition can be grouped to recognize the object. This is performed by noting the locations of the correlation peaks in relation to each other. A fuzzy score is attached to these relationships. A large fuzzy score indicates that features have been identified and are located in positions that are indicative of the target.

5. Summary

Foveation through a PCNN is quite possible. These foveation points resemble to some degree those measured in humans. Foveation points are generally close to the corners and edges of segments. The PCNN has the inherent ability to segment images and these segments can be easily filtered to provide foveation points. Such points are extremely useful for target recognition as shown here by example.

6. Acknowledgements

REFERENCES

- [Brasher,94] J. D. Brasher, J. M. Kinser, "Fractional-Power Synthetic Discriminant Functions", *Pattern Recognition* **27**(4), 577-585 (1994).
- [Eckhorn,90] R. Eckhorn, H.J. Reitboeck, M. Arndt, P. Dicke, "Feature Linking via Synchronization among Distributed Assemblies: Simulations of Results from Cat Visual Cortex," *Neural Computation*, **2**, 293-307, (1990).
- [Johnson,94] J. L. Johnson, "Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariances for images", *Appl. Opt.* **33**(26), 6239-6253 (1994)
- [Srinivasan,96] R. Srinivasan, J. Kinser, M. Schamschula, J. Shamir, H. J. Caulfield, "Optical Syntactic Pattern Recognition Using Fuzzy Scoring", *Optics Letters*, **21**(11) 815-817 (1996).
- [Yarbus,65] A. L. Yarbus, *The Role of Eye Movements in Vision Process*, USSR, Moscow, Nauka (1965).

FIGURES

1. A Typical Foveation Pattern.

2. The PCNN Neuron.
3. The Logical Flow of the Foveation System.
4. A Pulse Image and the Filtered Version of that Image.
5. Handwritten Characters and Their Foveation Points as Determined by the PCNN-Based Model.
6. The Foveation Points for the Smoothed Face Image for 10, 20, 30, 40 and 50 Iterations.
7. Foveation points Overlaid on a “Real-World” Image.
8. Typical Handwritten Characters.
9. Logical Flow of the Recognition System.
10. Barrel Distorted Images Centered on a Foveation Points.

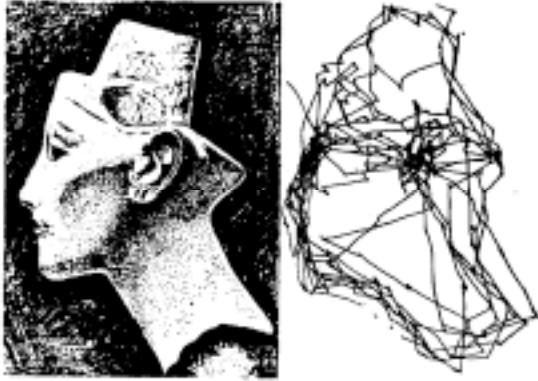


Fig. 1. A Typical Foveation Pattern. [Rybak,65]

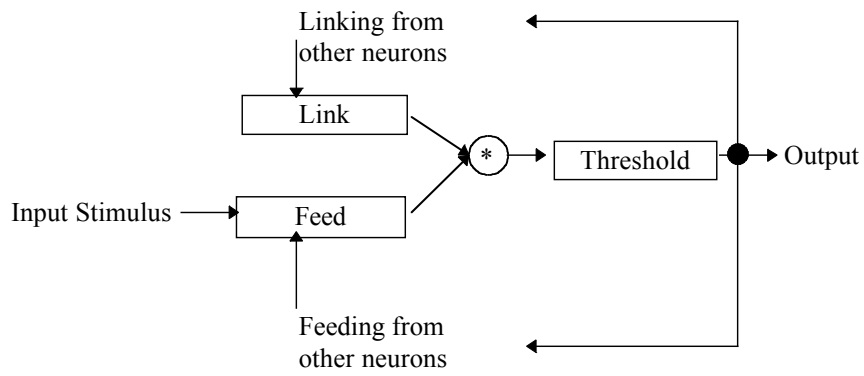


Fig. 2. The PCNN Neuron.

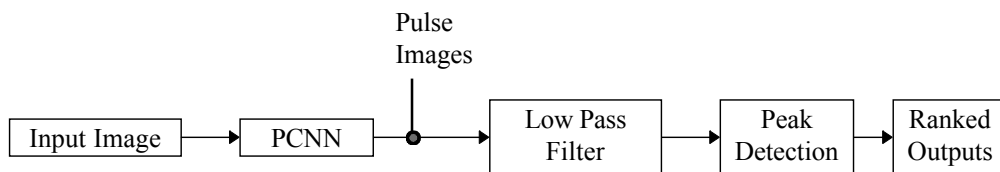


Fig. 3. The Logical Flow of the Foveation System.



Fig. 4. A Pulse Image and the Filtered Version of that Image.

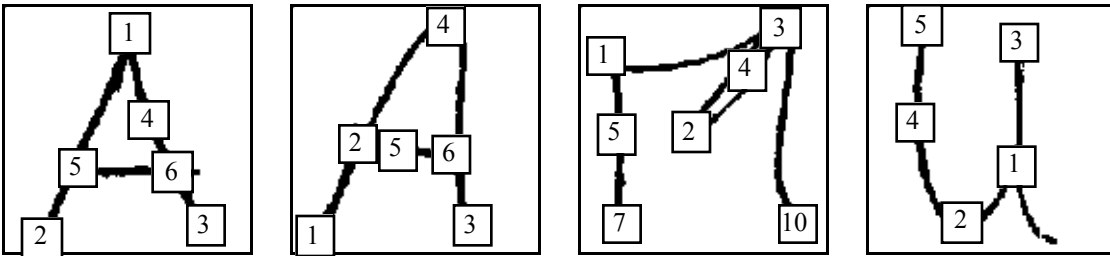


Fig. 5. Handwritten Characters and Their Foveation Points as Determined by the PCNN-Based Model.

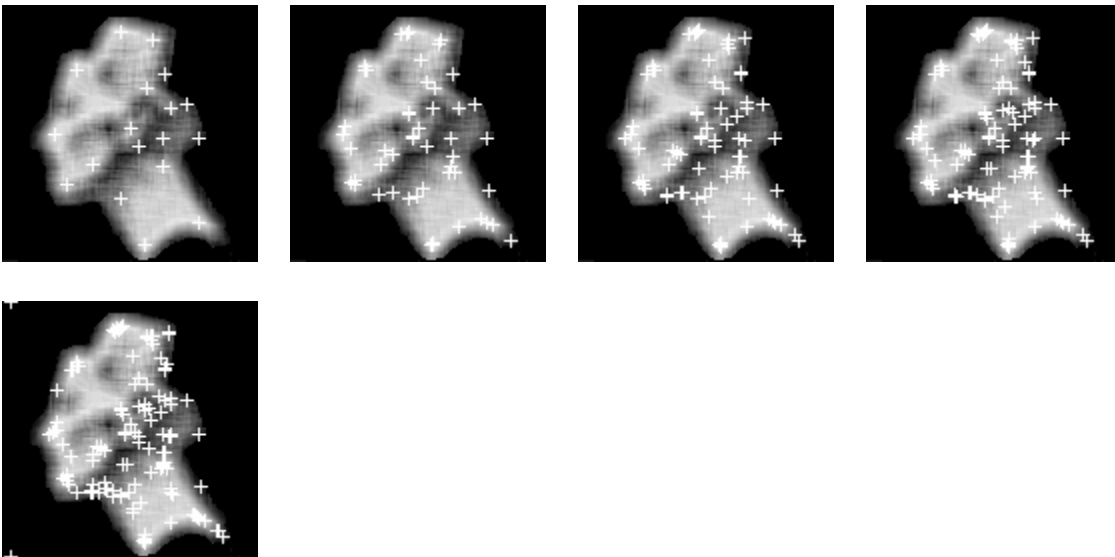


Fig. 6. The Foveation Points for the Smoothed Face Image for 10, 20, 30, 40 and 50 Iterations.

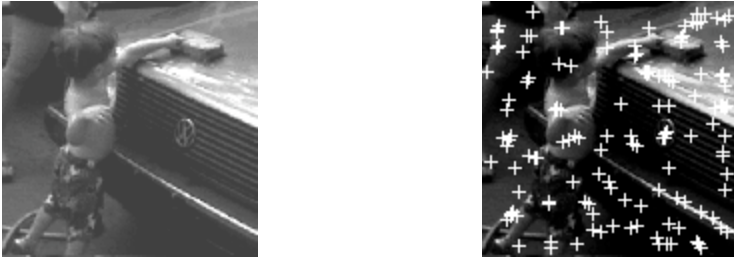


Fig. 7. Foveation points Overlaid on a “Real-World” Image.

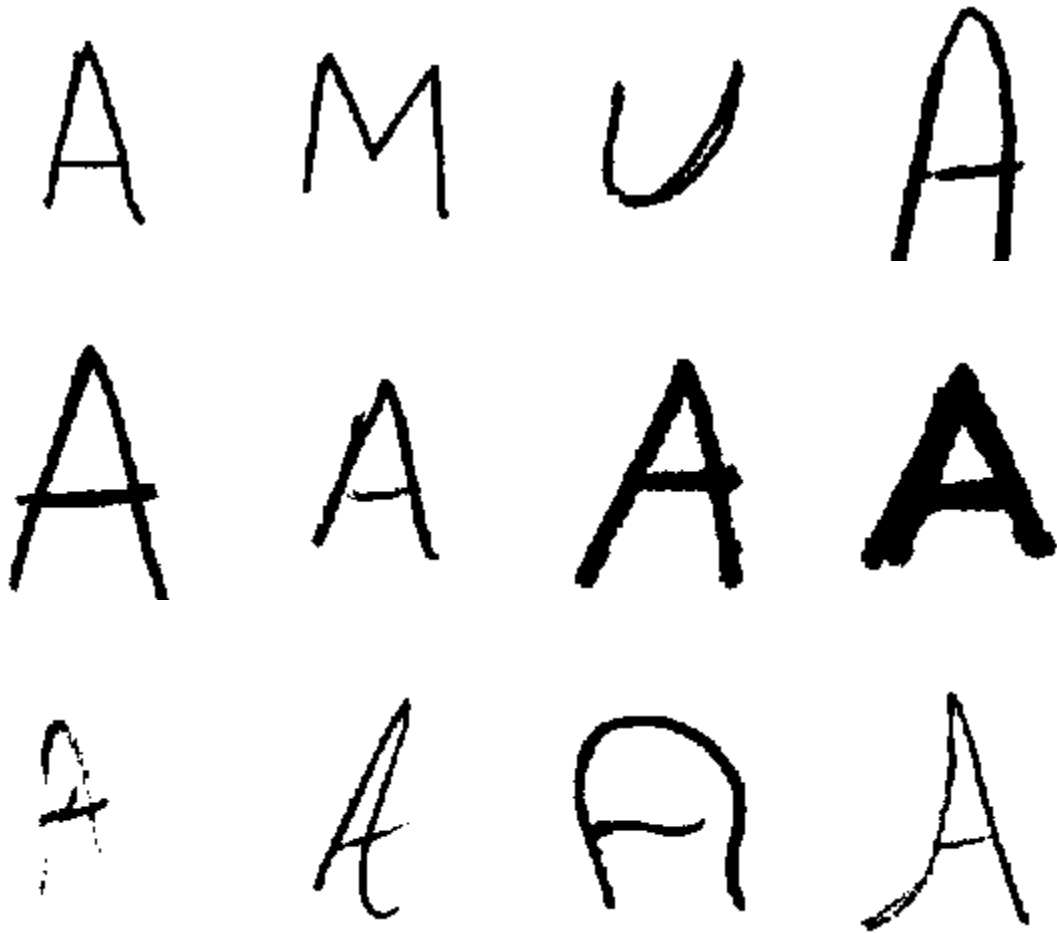


Fig. 8. Typical Handwritten Letters.

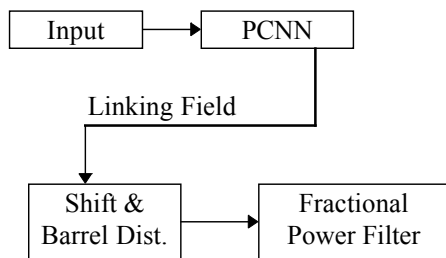


Fig. 9. Logical Flow of the Recognition System.

Table 1.

Category	Average	Low	High	Std. Dev.
Target	0.995	0.338	1.700	0.242
Non-Target	0.137	0.016	0.360	0.129