Shape-recognition using randomly selected pixel-pair neurons

Valerie Rose and Jeffrey Johnson
Department of Design and Innovation
The Open University
Milton Keynes, MK7 6AA, England
j.h.johnson@open.ac.uk and vrose@ukonline.co.uk

Abstract

Designing low level recogniser/filters for vision systems restricts their potential. We show how *retinal neurons* based on pixel-pair sensors can recognise shapes. The next step is evolving optimum retinal neurons to support the evolution of vision systems that can abstract their own higher level constructs.

1. Introduction

Our group is investigating a new vision architecture based on a multilayer representation [1].This architecture is intended to produce machine vision systems that:

- evolve appropriate retinal configurations
- evolve connectivities to represent spatial relationships
- abstract their own higher level constructs
- have levels which are integrated by new relational mathematics

A key feature of the architecture is a *multilevel representation*, with pixels at the lowest level, and objects and scenes at higher levels.

At the lowest layer, configurations of pixels respond to objects and are connected to an hypothetical neuron that fires when activated by appropriate inputs. We call these *retinal configurations*. The outputs of lower level neurons feeds forward through the system, ultimately allowing objects and scenes to be recognised.

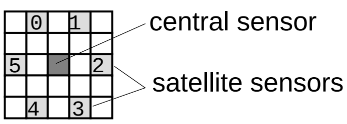

Figure 1. Inputs to a low-level neuron

Elsewhere [1] we have experimented with retinal configurations based on fixed pixel positions, as illustrated in Figure 1. Here there is a central sensor, surrounded by six satellite sensors. The idea is that if the central sensor detects darkness, the neuron is triggered. There are sixty four light-dark responses for the six satellite pixels, (Fig 2 We assume that there are sixty-four associated retina neurons. Once the central neuron is triggered, the neuron fires if the six satellite pixels match the black-white greyscales of the image. Thus, every dark pixel in an image responds to one of these sixty-four retinal neurons, according to the black-white states of its satellite pixels.

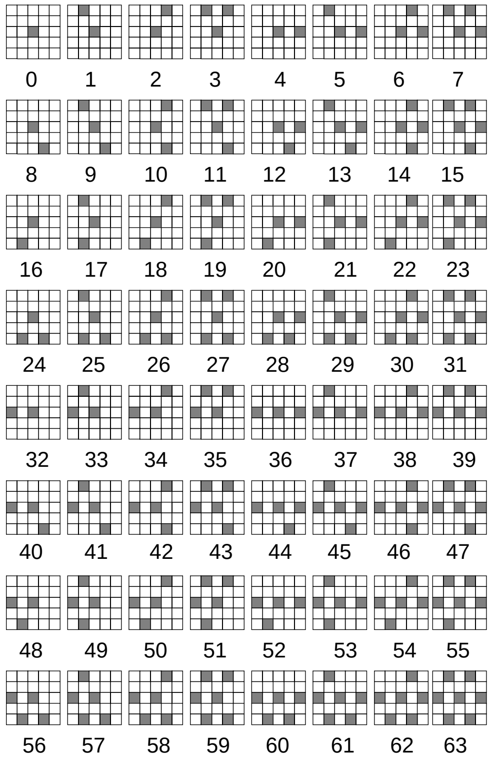

Figure 2. 64 retinal neuron patterns

Figure 3 shows the three shapes used in the experiments reported in this paper. Each shape responds differently to the retinal neurons. For example, the diamond shape has more type-57 responses than the circle or square, because this configuration of the satellite sensors responds well to the sloping edges. For a given shape, the numbers of responses for each retinal neuron can be counted, giving a 64-element *response vector* for each shape.

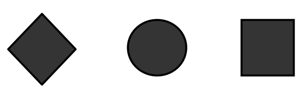

Figure 3. The simple shapes used for our experiments

It has been shown that these vectors contain sufficient information to discriminate the simple shapes, in the context of our hierarchical architecture, which is described in detail in [1].

2. Random Neural Pair Generation

An essential feature of our architecture is that it must be able to *adapt* to changes in objects and scenes. We believe that this pre-empts approaches to machine vision in which programmers *design* a fixed retinal processing architecture. In particular we believe that the approach typified by our designing the sixty four retinal neurons in Figure 2 inevitably leads to vision systems that will be limited in their recognition ability, and incapable of adapting to radically new objects and scenes.

We conclude from this that we need to *generate the retinal neuron configurations at random*. There are many questions associated with this, including:

- compared to one plus six sensors used in Figure 1, how many sensors should be connected together to produce a retinal neuron?
- what should be the maximum or minimum 'diameter' of the neuron?
- how many neurons are required for successful pattern recognition?

To begin our experiments we used pairs of pixel sensors connected to hypothetical neurons, as shown in Figure 4. They can occur at any distance from one another within the image, as shown in Figure 4.

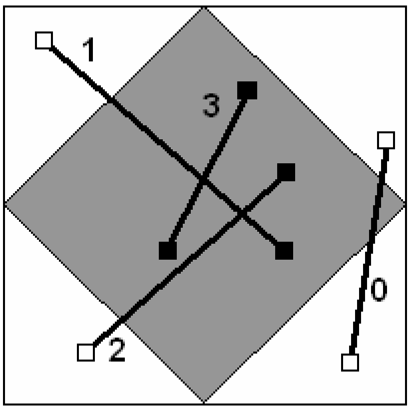

Figure 4. The four pixel-pair configurations
0 - background-background
1 - background-foreground
2 - foreground-background
3 - foreground-foreground

There are four possibilities for the pixel pair: background-background (0) background-foreground (1), 2 - foreground-background (2), and foreground-foreground (3).

For our experiments, we generated sixty pixel pair configurations. These random configuration were fixed and used for training and recognition. The retinal neurons are illustrated in Figure 5.

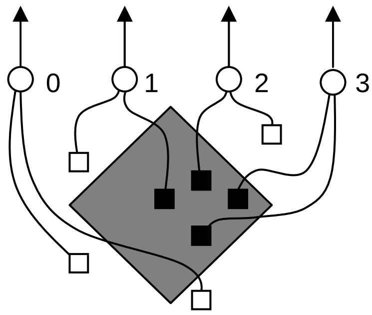

Figure 5. Pixel pair retinal neurons

Three 'template' shapes were generated using one shape of each class shown in Figure 3 as the training set. If the sensor black-white inputs correspond to the fixed black-white input configuration of the neuron, then that neuron fires. For example, if the neuron is of Type 0 and both input pixels are white, the neuron fires.

3. Recognising unseen shapes

For a given set of pixel pairs and a given shape, the neuron response is recorded as a template. These templates are matched against unseen shapes and used to recognise them. Thus each unseen shape gets a score of the number of neurons that were matched and fired. We apply a winner-takes all strategy, and the shape with the highest number of matches is recognised. If two or more shapes have the same score, a non-classification is made.

Circle	Diamond	Square
88/88	87/88	88/88

Table 1. Recognition of the three shapes

For this experiment we used eighty eight examples of each test shape. All the circles, all the squares, and all but one of the diamonds were correctly recognised (Table 1). This result strongly suggests that this approach of matching shapes against randomly generated retinal neurons is viable.

Trial	Circle	Diamond	Square	%
1	88/88	87/88 *	88/88	99%
2	88/88	88/88	88/88	100%
3	88/88	80/88 *	84/88 *	94%
4	88/88	85/88 *	88/88	98%
5	88/88	88/88	88/88	100%
6	88/88	88/88	88/88	100%
7	88/88	88/88	88/88	100%
8	88/88	88/88	88/88	100%
9	87/88 *	88/88	88/88	99%
10	88/88	88/88	88/88	100%

Table 2. Ten recognition trials (60 neurons)

To strengthen this conclusion the experiment was repeated ten times. The results of these experiments are shown in Table 2.

These results support the conclusion that the randomly generated retinal neurons can be used for shape recognition, though they show that errors can occur.

In order to see if these results could be improved, the experiment was repeated using sets of one hundred randomly generated neural pairs. The results in Table 3 suggest that adding extra neurons can improve the pattern recognition performance.

Trial	Circle	Diamond	Square	%
1	88/88	88/88 +	88/88	100%
2	88/88	88/88	88/88	100%
3	88/88	84/88 +	88/88 +	100%
4	88/88	87/88 +	88/88	99%
5	88/88	88/88	88/88	100%
6	88/88	88/88	88/88	100%
7	88/88	88/88	88/88	100%
8	88/88	88/88	88/88	100%
9	88/88 +	88/88	88/88	100%
10	88/88	88/88	88/88	100%

Table 3. Ten further trials with 100 neurons

Not surprisingly, the location as well as the type of configuration of the random pairs affects performance, as does the number of pairs selected.

4. Discrimination by configuration

The next set of experiments is concerned with the relative ‘usefulness’ of the different configurations.

Table 4 indicates that, as one would expect, considering the relatively large shapes we are using, that the largest proportion of the 60 pairs are of type ‘3’, black-black, for each of the three template shapes.

Trial	Circle	Diamond	Square
1	44	24	56
2	41	19	54
3	40	19	53
4	37	17	53
5	40	19	56
6	34	12	55
7	36	20	55
8	34	14	53
9	38	16	55
10	42	21	54

Table 4. Frequency of Configuration 3

Using just configuration 4 for matching gave an improvement in all the scores over the first set of experiments (Table 2), and no worsening of any score. The performance for diamonds is notably better.

So these black-black pairs are being generated in sufficient quantity to give reasonably good recognition.

Increasing the number of random pixel-pairs to 100 further improves performance

The next set of experiments restricted the permitted configurations to types ‘1’ and ‘2’, white-black and black-white.

This gives some slight improvements with recognition of diamonds in Trial sets 1and 3, but some deterioration in recognition with trial sets 3, 4, 5, 7, 9 and 10.

Another factor which is possibly affecting performance is that the ‘not 1 and 2’ configurations, ‘0’ and ‘3’, are opposites – background/background versus foreground/foreground, so that grouping them together into one type may reduce discriminatory ability. In other words, it seems that information about ‘edges’ and ‘not edges’ without further categorization of the ‘not edges’ as ‘shape’ or ‘background’ is insufficient for reliable recognition.

Again selecting 100 random pairs increases the quantity of type ‘1’ and ‘2’ configurations and performance improves correspondingly.

Increasing to 130 pairs brings further improvement – 100% recognition for all but the diamonds, with one trial giving a marked deterioration . So, in general, even when the number of pairs is increased producing more type ‘1’ and ‘2’ pairs, recognition is not as reliable as for all four configurations or types ‘3’ and ‘not 3’.

5. Distance between paired pixels

So far there has been no restriction on the separation of the pixel pairs. In this section we investigate whether restricting the distance affects the results.

It appears that restricting the permitted distance between the pixels in each pair adversely affects performance for a given number of random pairs.

The recognition experiments were repeated with the distance, in both horizontal and vertical directions being reduced to <=10 to <=5 and finally to <=3. For 60 pixel-pairs recognition deteriorates correspondingly.

Increasing to 150 pairs and setting the distance to <=10 pixels gives 100% recognition for all except the diamonds in one of ten trial sets. However, with 150 pairs and any distance, all shapes are correctly classified across all ten random sets, so restricting the distance appears to give no advantage.

One might expect that restricting the distance would provide useful 'local' discriminatory information. Possibly the reason this does not appear to do so is the relatively small number of pixel-pairs, and the lack of redundancy. When more random pairs are generated, there is a better chance of gathering enough useful information to compensate for this.

6. Informal Discussion

These experiments have addressed a number of questions. They have been concerned with how randomness might used as the basis of recognition and what gives useful information in order to develop strategies that able it to repeat successful behaviours.

In order to succeed at this, the system needs to be able to quantify the amount of information is getting from parameters such as distance between pixels, their location, the configurations they form in combination, and so on. It may even be useful to base its evaluations on information theory related to the inputs and outputs on the various connections in some form of hierarchical neural network.

However this raises issues such as how to put a value on the relative quantities of information being conveyed. For example, how would the system differentiate between the amount of information obtained from a pair of pixels of type '3' configuration 10 pixels apart and a type '2' configuration pair 60 pixels apart? The answer to this probably lies in having feedback of some sort about the varying strengths of neuronal responses to the different input patterns as the system learns. Rolls and Deco [2] have recognized the contribution information theory can make to understanding communications among individual neurons and networks of neurons within the brain.

7. Conclusions and the next step

The conclusions from the experiments reported here are that:

- randomly generated retinal neurons using pairs of pixel sensors can be used to recognise shapes
- increasing the number of neurons from 60 to 150 improved performance
- of the four types of neuron, Type-3 appears to give the most information, but this rather inconclusive
- restricting distances between sensor pairs did not appear to improve recognition performance, although the performance improved with the number of neurons.

These conclusions have to be set in the context of the small number of test shapes we have used, and their simple nature.

The results support our belief that we will be able to use randomly generated neurons for recognising shapes and features. In future experiments we intend to use evolutionary methods to select the most appropriate neurons for particular purposes. It is our intention to 'breed' new sets of retinal neurons from those that perform best for a given application.

The result that restricting distance does not appear to improve performance is particularly interesting. Combined with the result that increasing the number of neurons improves performance, this suggest that evolutionary principles can best select the most appropriate distances for any particular application, provided enough neurons are used.

References

[1] Johnson, J.H., Rose, V., 'Autonomous Evolutionary machine vision systems', AROB-10, Beppu, Japan, M. Sugisaka (ed).

[2] E. T. Rolls, G. Deco, 2002, *Computational Neuroscience of Vision*, Oxford.