# A Connectionist Approach for Color Image Segmentation

V.V.Vinod† Santanu Chaudhury†† Jayanta Mukherjee‡ S.Ghose†

†Dept. of Computer Sc. & Engg., Indian Institute of Technology, Kharagpur, India – 721302
‡Dept. of Electronics & Elect. Commn. Engg., Indian Institute of Technology, Kharagpur, India – 721302
††Dept. of Electrical Engg., Indian Institute of Technology, Delhi, India – 110016

*Abstract* *In this paper a connectionist clustering strategy is presented for segmenting color images. First the local peaks in the 3-D R,G,B histogram are located. Then using these as the prototypes other patterns are classified to one of them. The prototype selection and classification networks have been analyzed. The prototype selection method employs only neuronal dynamics and therefore is faster than existing clustering neural networks. The classification network takes into account the distribution of the data and hence is less prone to misclassifications. Experimental results obtained by applying the network for segmenting one color image is presented.*

**Keywords** Segmentation, Clustering, Mode Separation, Neural Networks.

## I INTRODUCTION

Color segmentation is the process of grouping together image pixels based on their color values. Earlier works employed statistical methods for color segmentation using prior knowledge. However, in situations where the statistical properties are not known a priori, clustering techniques have to be employed. Most of the traditional clustering algorithms require the number of clusters and/or the clustering criterion to be specified beforehand. This information is rarely available a priori. Histogram based mode separation techniques have the capability to autonomously detect the number of clusters and would apply well in such situations. In this paper we propose a neural network strategy for color segmentation by a process similar to mode separation.

A number of clustering techniques based on histogram analysis have been proposed for color segmentation [1, 2, 3]. However, it has been shown that known techniques for modal analysis of histograms face difficulties when the histogram is characterized by extremely unequal peaks and broad or flat valleys [4]. In the case of color images the histogram is 3-dimensional and as a result modal analysis techniques are faced by a number of extra difficulties. The valley detection methods employed for clustering 1-D histograms, either do not become applicable or become computationally very expensive. Therefore, in general, the projections of the histogram onto lower dimensional spaces are considered and the modes so obtained are then combined using some criterion. For example in [1] the clusters are detected by finding the modes in the coordinate projections of the $L^*,H^o,C^*$ histogram and then separating them using the Fischer linear discriminant. The algorithm depends heavily on the presence of decisive modes in the 1-D histograms. In the absence of such modes the method selects a new set of features and proceeds. However, the feature selection process is computationally very intensive and results in a large amount of overall computation time. All 3-D clustering methods based on lower dimensional projections are faced with this drawback. It may be observed that, though modes may be absent in the 1-D projections, they will be present in the 3-D histogram. Mode separation in the 3-D histogram itself will be more direct and efficient. Recently Jolion *et al.* [5] have proposed a minimum volume ellipsoid based clustering technique. This method, however, does not work on the 3-D histogram, but on a set of non continuous points and extracts the clusters sequentially.

Connectionist clustering methods such as self-organizing feature maps [6] and ART models [7] provide alternate clustering strategies [8]. These models, except ART, require the number of clusters to be specified beforehand. And all these networks employ weight learning and is therefore slow. With respect to the histogram the prototypes of the clusters will be the local peaks. Fast neural networks employing only neuronal dynamics have been proposed for detecting multiple local peaks [9, 10]. We propose to detect the prototypes in the R,G,B histogram by such a network. A brief outline of the network is given in section II.

Another aspect of clustering is the process of classifying the patterns to one of the prototypes. Existing neural networks classifying patterns to the nearest prototype. Consequently contiguous regions may be split into different clusters. It would be desirable to maintain clusters as contiguous regions. Nearest neighbour classification will result in large errors when the clusters are of varying sizes and shapes. In order to avoid such difficulties an efficient classification strategy would have to take into account the histogram distribution in addition to the distance. We propose a classification network which takes into account these aspects. The detailed design and analysis of the network is given in section III. Experimental results obtained by segmenting the image of a natural object is presented in section IV.

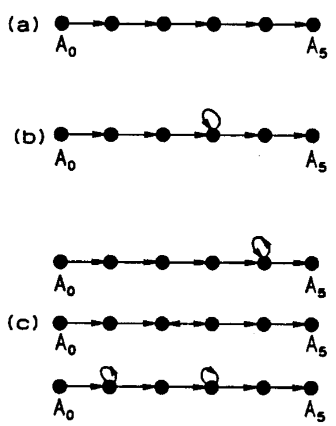



Figure 1: Schematic propagation of activations in a 1-D classification network

The parameters $c_1$ and $c_2$ determine the tradeoff between nearest neighbour classification and strictly following the frequency distribution. A higher (lower) value of $c_1$ will associate a higher (lower) weightage to the distance between the prototype and the pattern. A higher (lower) value of $c_2$ will associate a higher (lower) weightage to the distribution between the prototype and the pattern. For well-formed modes, where the number of cells with low frequency counts are less and placed apart, a lower value of $c_1$ will lead to better results. On the other hand, if the modes have large number of closely spaced zero valued cells in them, then a larger value of $c_1$ is suggested.

### *B Analysis of the Network*

The accumulator layer implements the process of evaluating the classification criterion. We focus on an accumulator neuron corresponding to a prototype and briefly discuss the properties of the classification network. In the context of the discretized pattern space represented by the histogram, two cells are considered to be in a contiguous region if there exists a path between them through other cells each with a nonzero frequency count. Such a path is essentially a path in the network through which the activation propagates from one neuron to another. The number of neurons in a path forms the length of the path. We analyze the factors considered by the network while classifying a given pattern. The study is presented with respect to a one dimensional histogram so as to keep the discussion brief.

Figure 1 schematically shows the propagation of activations along a path of length 5 from a neuron $A_0$ to a prototype neuron $A_5$. Since connections among accumulator neurons are only to immediate neighbours, this is the shortest path from $A_0$ to $A_5$. No activation from $A_0$ will reach $A_5$ until the fifth iteration. Let $w_i$ denote the quantity $c_1 + c_2 h_i$.

Figure 1(a) shows the activation reaching $A_5$ at $t = 5$.

$$\alpha_5^5 = w_1.w_2.w_3.w_4.w_5 = \prod_{i=1}^{5} w_i$$

This expression may be generalized for any path of length $l$ as

$$\alpha_l^l = \prod_{i=1}^{l} w_i$$

It may be seen that in the ideal case where every pattern is contiguous to exactly one prototype and $c_1 = 0$, classification can be done using this value. The product term will be nonzero only for that prototype with which the pattern is contiguous. However, due to nonidealities, $c_1 \neq 0$ and therefore the product term may have low nonzero activation for more than one prototype. The product will be of significant value only if all the $w_i$ are of high value, i.e., all cells in the path have high frequency counts. If the activation of the prototype is sufficiently high then classification is done. Otherwise, further properties of this path (and other paths) will have to be utilized. In the next iteration the network considers all paths of length 6.

The activation along the path from $A_0$ to $A_5$ in the $6^{th}$ iteration will include that shown in figure 1(a) and all paths of the form shown in figure 1(b). The former is the result of clamping the activation of neuron $A_0$. The latter is the contribution of paths of length 6. There will be four such paths. The total activation along all these paths is obtained as

$$\alpha_5^6 = \prod_{i=1}^{5} w_i (1 + \sum_{i=1}^{4} w_i)$$

In the above expression, the summation is the new term introduced. As a result, the earlier measure is now weighted with the total frequency count along the path. It may be observed that the summation will have a low value if there are more valleys and a higher value otherwise. This term may be viewed as counting the number of valleys.

The activations added to the above in the $7^{th}$ iteration are shown in figure 1(c). Including these $\alpha_5^7$ is obtained as:

$$\begin{aligned}\alpha_5^7 \;&=\; \prod_{i=1}^{5} w_i (1 + \sum_{i=1}^{4} w_i + \sum_{i=1}^{4} w_i^2 + \\ &\sum_{i=1, j>i+1}^{3} w_i w_j + \sum_{i=1}^{3} 2 w_i w_{i+1})\end{aligned}$$

## II PROTOTYPE SELECTION

The prototype selection network has two layers of neurons called accumulator layer and decision layer. Each layer has one neuron corresponding to each histogram cell. Interconnections are present among neighbouring neurons in the accumulator layer. Each neuron in the accumulator layer is connected to the corresponding neuron in the decision layer. The decision layer decides a prototype and inhibits other neurons in the neighbourhood of the prototype. For detecting the prototypes, the histogram values are given as inputs to the accumulator layer neuron and the network is updated synchronously. The network detects the prototypes by adaptive smoothing and adaptive thresholding. That is, the smoothing kernels and thresholds applied for detecting prototypes depend on the shape and size of the mode. Consequently the performance is not degraded by the presence of modes of widely varying sizes and shapes. The network is also capable of detecting prototypes in large regions with uniform distribution. The detailed architecture and analysis of the network may be found in [10, 11].

## III THE CLASSIFICATION NETWORK

The classification network consists of two layers, each layer having one neuron corresponding to each histogram cell. One of the layers is visible and the other hidden. The visible layer shall be referred to as the accumulator layer and the hidden layer shall be called the decision layer. Let $A_{ijk}$ and $D_{ijk}$ respectively denote the accumulator layer neuron and decision layer neuron corresponding to the cell $(i,j,k)$ of the histogram. The activation of $A_{ijk}$ after the $t^{\text{th}}$ synchronous iteration of the network shall be denoted by $\alpha_{ijk}^{t}$. Also, in subsequent discussions, neurons in the neighbourhood of $A_{ijk}$ shall mean the neurons corresponding to cells neighbouring $(i,j,k)$. The histogram count of cell $(i,j,k)$ shall be denoted by $h_{ijk}$.

### A The Network Architecture

When presented with a pattern to be classified, the network should try to find the closest prototype which can gather sufficient strength from the pattern. The network starts with minimal length paths and iteratively increases the length. For this each neuron is connected to neurons in its immediate neighbourhood. Let $\mathsf{N}_{ijk}$ denote the neurons in the immediate neighbourhood of $A_{ijk}$.

$$\begin{aligned}\mathsf{N}_{ijk} &= \{(l,m,n) \mid l \neq i, j \neq m, k \neq n \\ &\quad \text{and } |l-i| = 1, |m-j| = 1, |n-k| = 1\}\end{aligned}$$

The activations received by a neuron are proportional to its frequency count. This ensures that propagation across cells with lower frequency counts are less compared to that across cells with larger frequency counts. However, in order to allow classification of patterns not contiguous with any prototype a small positive constant is added to the frequency values of the cells. It may be observed that we are interested in paths starting from a given pattern and ending at the prototype. Hence propagation across a prototype is not needed. The interconnection weights from $A_{ijk}$ to $A_{lmn}$ may now be written as

$$W(A_{ijk}, A_{lmn}) = \begin{cases} c_1 + c_2 h_{lmn} & \text{if } (i,j,k) \in \mathsf{N}_{lmn} \\ & \text{and } A_{ijk}\ A_{lmn} \text{ are not prototypes} \\ c_1 + c_2 h_{ijk} & \text{if } i = l \text{ and } j = m \text{ and } k = n \\ & \text{and } A_{ijk}\ A_{lmn} \text{ are not prototypes} \\ 1 & \text{if } A_{lmn} \text{ is a prototype} \\ 0 & \text{otherwise} \end{cases}$$

where $c_1$ and $c_2$ are constants such that $c_1 + c_2 h_{lmn} < 1.0$. This restriction is in order to ensure that the activation decreases as the distance of the prototype from the pattern increases.

The activation function of the accumulator neurons is chosen as

$$f_a(x) = \begin{cases} 1 & \text{if } x > 1 \\ -1 & \text{if } x < -1 \\ x & \text{otherwise} \end{cases}$$

With the above interconnections, every prototype starting from the one nearest to the given pattern will iteratively accumulate activation from its neighbourhood. The decision layer accepts an accumulator layer neuron gaining sufficient activation as the correct prototype and inhibits others. For this purpose, only those decision layer neurons which correspond to the prototypes are relevant. Such decision layer neurons receive unit weighted interconnections from the corresponding accumulator layer neuron. The threshold $\phi$ of the decision layer neurons are fixed at a level close to the maximum value attainable by an accumulator layer neuron. The connections from the decision layer neurons (corresponding to the prototypes) to the accumulator layer neurons are set up as

$$W(D_{ijk}, A_{lmn}) = \begin{cases} -\infty & \text{if } i \neq l \text{ or } j \neq m \text{ or } k \neq n \\ \infty & \text{otherwise} \end{cases} \tag{1}$$

where $\infty$ stands for a large positive number. A high gain sigmoid function $f_d(\cdot)$ is used as the activation function of the decision layer neurons. When the activation of an accumulator neuron corresponding to a prototype exceeds the threshold, the corresponding decision layer neuron will get an activation near 1. The pattern is then classified to this prototype.

Consider a pattern mapping onto the neuron $A_{ijk}$. For classifying this pattern the neuron $A_{ijk}$ is clamped at +1.0 and the network is updated synchronously. Each neuron $A_{lmn} \in \mathsf{N}_{ijk}$ will receive activation proportional to $c_1 + c_2 h_{lmn}$. By iteratively accumulating the activations, the neurons will increase in activation. When a prototype acquires activation greater than the threshold, the corresponding decision layer neuron becomes active. The high positive weight from the decision layer will sustain the winner accumulator neuron. The large negative weight will cause all other neurons to be strongly inhibited. The pattern is then classified with this prototype.

In this expression a number of extra terms are introduced. The terms $w_i w_j$ for $j > i$ are essentially detecting the presence of two valleys. It may be noted that the term $\sum w_i$ may be of same value due to the presence of one deep valley or two valleys of half the depth. In this iteration the network is now explicitly looking for two distinct valleys. The effect of the last summation is to detect adjacent cells with low frequency counts. In other words, the number of valleys of width 2 are being used this time to differentiate the two prototypes. It may be noted that this term is given a higher weightage than the previous ones. Hence the presence (absence) of adjacent valleys are more important for the network than the presence (absence) of non adjacent valleys. The other term introduced is $\sum w_i^2$, which is augmenting the effect of $\sum w_i$. However, since all $w_i < 1$, the former will have a low value compared to the latter and may be ignored.

The expression for the activation along the path at the $8^{th}$ iteration may be written as

$$\begin{aligned}\alpha_5^8 &= \alpha_5^7 + \prod_{i=1}^{5} w_i \Big(\sum_{i=1}^{4} w_i^3 + \sum_{i=1,j>i}^{3} w_i w_j^2 + \\ &\quad \sum_{i=1,j>i}^{3} w_i^2 w_j + \sum_{i=1,j>i,k>j}^{3} w_i w_j w_k\Big)\end{aligned}$$

The term $\sum w_i w_j^2$ may be ignored compared to the term $\sum w_i w_j$ present in the expression for $\alpha_5^7$. It only serves to slightly add on to the effect of the latter term. The quality of the path being considered is the same, namely presence (absence) of two distinct valleys. On similar grounds, the first and the third summations also can be ignored. The third term is the product of 3 distinct frequency values. As in the previous case, the network is looking for the presence of 3 valleys, consecutive as well as disjoint.

Another factor playing an important role in the classification scheme is the distance. Since the network considers paths of increasing sizes, this factor is automatically accounted for while classifying. For example, in the $8^{th}$ iteration, the activation accumulated at $A_8$ from $A_0$ will consist of the product term only. Whereas, for $A_5$ the product term has been enhanced by the various factors discussed above. Since all these factors are positive, the activation accumulated along a path increase with the number of iterations. It is possible to strike a balance between the distance and the other measures discussed above. This factor may be controlled by varying $c_1$ for small effects and varying the neighbourhood for gross effects. In general, for a prototype at a distance $l$ from the pattern, we see that in the $(t+l)^{th}$ iteration, the network considers

- the valleys of width up to $t$
- the presence/absence of $t$ distinct valleys
- the contiguity of the path as a whole
- the distance between the pattern and prototype

The above discussion has been done with respect to a 1-D histogram. In the case of the three dimensional R,G,B histogram, in addition to the width and depth the area and volume of the valleys will also be taken into account.

## IV Results

In this section we present the results obtained by applying the proposed strategy to segment the image of a natural object. The 3-D $R,G,B$, histogram of the image was used as input to the network. The range of each color component (0 to 255) was divided into 25 equal sized cells for constructing the histogram. The histogram values were scaled to lie in the range [0.0,0.4] (i.e. $h = 0.4$). This scaling was dictated by the prototype selection network.

The activation function $f_d(\cdot)$ was chosen as:

$$f_d(x) = \frac{1}{1 + \exp(-40x)}$$

A discussion regarding the choice of parameters for the prototype selection network may be found in [11]. The classification network was employed with $c_1 = 0.5$ and $c_2 = 1.25$. The threshold of decision layer neurons was chosen as 0.9.

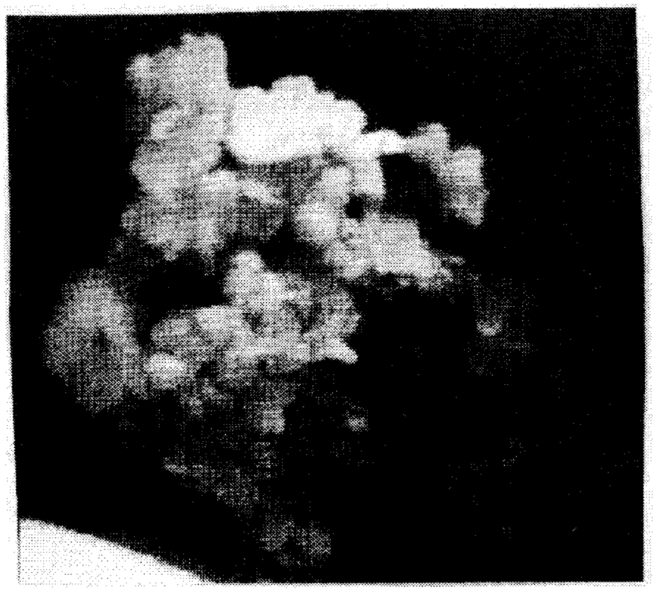
Figure 2(a) shows the 256 × 256 image of a flower vase with flowers. The segmented image is give in 2(b). The network correctly detected all the different colored regions of the image. The original image had smooth variations of colors resulting in spread out modes without clear demarcations. In spite of this, the network correctly segmented the image. The number of iterations required for detecting the prototypes was 200. The average number of iterations required for classifying the image pixels was 4.7 per histogram cell with nonzero frequency count.

## V Conclusion

In this paper we have presented a connectionist clustering strategy for color image segmentation. The clusters are extracted directly from the R,G,B histogram. Consequently the method works better than other strategies employing lower dimensional projections of the histograms. The prototype selection method employs only neuronal dynamics and therefore is faster than existing clustering neural networks. The proposed approach classifies patterns based on the distribution of the data unlike other clustering neural networks which perform nearest neighbour classification. Therefore it is less prone to misclassifications.

## References

[1] M. Celenk, "A color clustering technique for image segmentation," *Computer Vision, Graphics and Image Processing*, vol. 52, no. 2, 1990.



(a)

(b)

Figure 2: Flower vase. (a) original image. (b) segmented image


[2] A. Khotanzad and A. Bouarfa, "Image segmentation by a parallel, non-parametric histogram based clustering algorithm," *Pattern Recognition*, vol. 23, no. 9, pp. 961–973, 1990.

[3] Y. W. Lim and S. U. Lee, "On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques," *Pattern Recognition*, vol. 23, no. 9, pp. 935–952, 1990.

[4] P. K. Sahoo, S. Soltani, K. C. Wong, and W. C. Chen, "A survey of thresholding techniques," *Computer Vision, Graphics and Image Processing*, no. 41, pp. 233–260, 1988.

[5] J. M. Jolion, P. Meer, and S. Bataouche, "Robust clustering with applications in computer vision," *IEEE Transactions on Pattern Analalysis and Machine Intelligence*, vol. 13, no. 8, pp. 791–802, 1991.

[6] T. Kohonen, *Self-Organizing and Associative Memory*. Springer-Verlag, 1984.

[7] S. Grossberg, "Nonlinear neural networks: Principles, mechanisms and architecture," *Neural Networks*, vol. 1, pp. 17–61, 1988.

[8] Y. Pao, *Adaptive Pattern Recognition and Neural Networks*. Addison-Wesley, 1989.

[9] V. V. Vinod, S. Chaudhury, J. Mukherjee, and S. Ghose, "A connectionist approach for gray level image segmentation," in *Proceedings of the* $11^{th}$ *International Conference on Pattern Recognition, Delft*, 1992. In Press.

[10] V. V. Vinod, S. Chaudhury, S. Ghose, and J. Mukherjee, "A connectionist approach for peak detection in Hough space," *Pattern Recognition*. In Press.

[11] V. V. Vinod, S. Chaudhury, J. Mukherjee, and S. Ghose, "A connectionist approach for clustering with applications in image analysis," *IEEE Transactions on Systems Man and Cybernetics*. In Press.