

# Chapter 15

## A New Vision Inspired Clustering Approach

Dequan Jin and Zhili Huang

**Abstract** In this paper, a new clustering approach by simulating human vision process is presented. Human is good at detecting and segmenting objects from the background, even when these objects have not been seen before, which are clustering activities in fact. Since human vision shows good potential in clustering, it inspires us that reproducing the mechanism of human vision may be a good way of data clustering. Following this idea, we present a new clustering approach by reproducing the three functional levels of human vision. Numeric examples show that our approach is feasible, computationally stable, suitable to discover arbitrarily shaped clusters, and insensitive to noises.

**Keywords** Neural field theory · Clustering Analysis · Amari's model · Stationary solution · Dynamical system

### 15.1 Introduction

Clustering is primitive in human learning activities [1]. All the unsupervised learning activities can be considered as clustering activities [2]. The most fundamental human clustering activity may be in human vision, for instant, in object

---

D. Jin (✉)  
School of Mathematics and Information Science, Guangxi University,  
No. 100 Daxue Road, Nanning, Guangxi, China  
e-mail: dqjin@yahoo.cn

D. Jin  
School of Mathematics and Statistics, Xi'an Jiaotong University,  
No. 28 West Xianning Road, Xi'an, Shaanxi, China

Z. Huang  
School of Mechanical Engineering, Guangxi University,  
No. 100 Daxue Road, Nanning, Guangxi, China

detection or segmentation, which are elementary visual activities. Human vision is highly complex. A simple visual activity may involve a lot of neuronal structures. If we just investigate one or several ones of these structures, it may be not enough to illustrate the visual activity. But if we consider the whole visual system that includes most of these coupled structures, it may become too complicated to be analyzed. Instead of considering individual neuronal structure, some investigations on visual perception indicates a feasible way to avoid these difficulties, which reproduce the functional levels of human vision, each of which might be structurally distributed [3, 4].

In this paper, a new clustering approach is presented, by considering data as an image and clustering them by three-level neural field system for visual perception. Method for determining the range of excited regions in the activation distribution of neural field is also introduced.

## 15.2 Levels for Human Vision

Generally speaking, the whole procedure of human vision contains three functional levels at least, including the transfer level, the planning level, and the motor control level.

### 15.2.1 Transfer Level

In this level, eyes and some low-level neuronal structures are involved. Eyes act as sensors, whose main task is to accept the light intensity distribution and transform them into neural signals distributions. The visual information may be subjected to some spatial and temporal transformations in this level induced by retina and some neuronal structures [4].

Suppose the objects are static in visual field. Then the light intensity distribution is usually presented as a static image that consists of  $N$  light points  $\{x_i\}_{i=1}^N$ , which can be described as [5]

$$I(z) = \frac{1}{N} \sum_{i=1}^N \delta(z - x_i) \quad (15.1)$$

where

$$\delta(z - x_i) = \lim_{\sigma \rightarrow 0} g(z, \sigma) \quad (15.2)$$

$g(z, \sigma)$  is a Gaussian function

$$g(z, \sigma) = \exp\left(-\frac{\|z\|^2}{\sigma^2}\right) \quad (15.3)$$

The visual information  $I(z)$  may be subject to some filtering effects in the transfer level, the most significant one of which in object detection and segmentation is the blurring effect induced by the retina, which is usually described as the Gaussian filtering process. Then the output neural signal distribution  $S(z)$  is given by

$$S(z) = S(z, \sigma_{in}) = \int_{\Omega} I(z - z')g(z, \sigma_{in})dz' \quad (15.4)$$

$\sigma_{in}$  is a scale parameter that can be understood as the distance between object and eye, or the curvature of crystalline lens [6]. If light points are too close that under the resolution of eyes, they cannot be identified in  $S(z)$  individually.  $S(z)$  is the output of the transfer level, as well as the input of the planning level.

### 15.2.2 Planning Level

In this level, the input neural signal  $S(z)$  would be processed by neurons. A visual perception is presented in the form of activation distribution of these neurons.

There are many neural models in describing the activity of neurons. A popular one of them is the Amari's dynamical neural field [7]:

$$\tau \dot{u}(z, t) = -u(z, t) + \int_{\Omega} w(z, z')\theta(u(z', t))dz' + S(z) - h \quad (15.5)$$

The vector space  $\Omega$  is called perceive space.  $\tau$  is a positive time constant.  $h$  is the resting level parameter. The region  $\{z \in \Omega : u(z, t) > 0\}$  is called excited region, denoting the activated neurons. The excited region usually corresponds to a perceived pattern.  $\theta(u)$  is a monotonically increasing nonlinear threshold function satisfying that  $\lim_{u \rightarrow -\infty} \theta(u) = 0$  and  $\lim_{u \rightarrow +\infty} \theta(u) = 1$ , for instant, the step function. It describes the neural field feedback of each excited point to its neighboring positions in  $\Omega$  with an interaction strength which is determined by interaction function  $w(z, z')$ .

Mostly, the interaction function  $w(z, z')$  is isotropic and usually written as  $w(z - z')$ . In this case,  $w(z)$  is also called as the interaction kernel of the neural field. Approximating the neurophysiologic lateral interaction among neurons, the lateral interaction of neural field is usually assumed to be locally exciting and globally inhibiting. One of the typical interaction kernels is the difference of Gaussian (DoG) functions with constant inhibition, given by

$$w(z) = Ag(z, \sigma) - Bg(z, \gamma\sigma) - h_{ker} \quad (15.6)$$

where  $\gamma > 1$ .

There are three important types of stable solutions to system (15.5), which are  $\phi$ -solution, “bubble”-solution and  $\infty$ -solution:

1. An equilibrium solution  $u^*(z)$  is called  $\phi$ -solution if  $u^*(z) \leq 0$  for all  $z \in \Omega$ ;
2. An equilibrium solution  $u^*(z)$  is called “bubble”-solution if  $u^*(z) > 0$  for  $z$  in a subset  $D \subset \Omega$ ;
3. An equilibrium solution  $u^*(z)$  is called  $\infty$ -solution if  $u^*(z) > 0$  for all  $z \in \Omega$ .

### 15.2.3 Motor Control Level

Motor control level aims at sending out control signals to specific organs, for instant, eyes, according to the neuron activity  $u(z, t)$  which is the output the planning level. For different purposes, this level would have different descriptions. For instant, in the investigation on saccadic motor planning, to control eyes to stare at an object in visual field, let  $z^* = \frac{P}{M}$ , where  $P = \int_{\mathbb{R}} z\theta(u^*(z))dz$  and  $M = \int_{\mathbb{R}} \theta(u^*(z))dz$ , then  $z^*$  is the density center of activation distribution which corresponds to the center of object [4].

## 15.3 Clustering Approach Based on Vision

Since human vision shows good potential in clustering, it is possible to find a feasible clustering approach by simulating visual mechanisms. In this section, we present a new clustering approach by reproducing the three levels of vision. Some numeric examples are given to show the feasibility and advantages of our approach.

### 15.3.1 Transfer Level for Clustering

In this level, the first thing is to transform data set  $X = \{x_i \in \mathbb{R}^n : i = 1, 2, \dots, N\}$ , to image which can be accepted by visual system by

$$I(z) = \frac{1}{N} \sum_{i=1}^N \delta(z - x_i) \quad (15.7)$$

In this way, we obtain a data distribution  $I(z)$ .

The data distribution  $I(z)$  is transformed into neural input distribution  $S(z)$  by a Gaussian filtering process:

$$S(z) = S(z, \sigma_{in}) = \int_{\Omega} I(z - z')g(z, \sigma_{in})dz' \quad (15.8)$$

### 15.3.2 Planning Level for Clustering

The aim of this level is to discover clusters in data, i.e., produce a perception of clusters, according to the neural signal distribution  $S(z)$ .

Since Amari's model achieves successes in illustrating phenomena in visual perception, we also employ it in the planning level for clustering, which is given by (15.5):

$$\tau \dot{u}(z, t) = -u(z, t) + \int_{\Omega} w(z, z')\theta(u(z', t))dz' + S(z) - h \quad (15.9)$$

where  $\Omega \subset \mathbb{R}^n$ .

As soon as  $S(z)$  being transferred to the planning level, the neural field begins to evolve, until the field reaches its steady state  $u^*(z)$ . Several bubbles, i.e., excited regions may be sustained in  $u^*(z)$ , whose number and range generally depends on the input  $S(z)$ , the kernel  $w(z)$  and the resting level  $h$  of the neural field. By grouping the data located in the same connected excited region into a cluster, the clusters of the data set  $X$  are perceived.

### 15.3.3 Motor Control Level for Clustering

In visual perception, the motor control level sends out control signal based on the activation distribution  $u^*(z)$  given by the planning level, so that people can react based on their perception in the planning level, corresponding to outside visual stimulus. In the motor control level, we introduce some methods to point out the range of connected excited regions in  $u^*(z)$ .

When a connected excited region in  $u^*(z)$  is convex, its range is equal to the attraction domain of a corresponding equilibrium point of the gradient dynamical system

$$\frac{dz}{dt} = \nabla u^*(z) \quad (15.10)$$

By estimating its corresponding attraction domain, we can estimate the range of a connected excited region. A feasible way for estimating the domain of attraction of such a system is presented in [8], which employs an iterative expansion approach. Details can be seen in [8].

## 15.4 Algorithm and Examples

On the basis of the above strategies, for a data set  $X = \{x_i \in \mathbb{R}^n : i = 1, 2, \dots, N\}$ , we present a clustering approach as following:

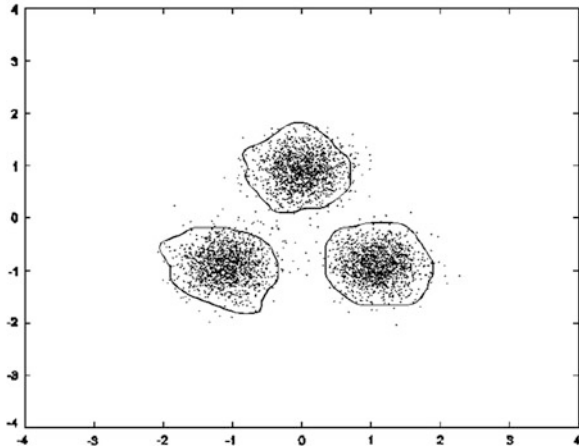
1. Select a scale  $\sigma_{in} > 0$ , and the interaction kernel  $w(z)$ . Compute the signal distribution  $S(z)$  by (15.8);
2. Let  $u(z, 0) = -h$  for  $z \in \Omega$ . Compute the steady state  $u^*(z)$  of system (15.9). If there are  $m$  excited regions, take all the data points in the same connected excited region into a cluster, denoting by  $C_j$ ,  $j = 1, 2, \dots, m$ . If there are unlabeled data points, go to step 3; else, let  $M = m$ , go to step 4.
3. If the unlabeled data points locate in excited regions, then group them to the clusters corresponding to these excited region; else, if the unlabeled data points locate in some peaks with negative activation, group the data locating in the same peaks into new clusters  $C_{m+i}$ ,  $i = 1, 2, \dots, \tilde{m}$ . Let  $M = m + \tilde{m}$ .
4. Let  $Sc = \{C_j\}_{j=1}^M$ , then  $Sc$  is the clustering result.

To show the feasibility of our approach, we give some numeric examples as shown in Figs. 15.1 and 15.2. The kernel  $w(z)$  is given by (15.6). Parameters are given as  $\sigma = 0.07$ ,  $\gamma = 1.1$ ,  $\tau = 0.1$ ,  $A = 1.2$  and  $B = 0.1$ ,  $h_{ker} = 0.002$ ,  $h = 0.02$ .

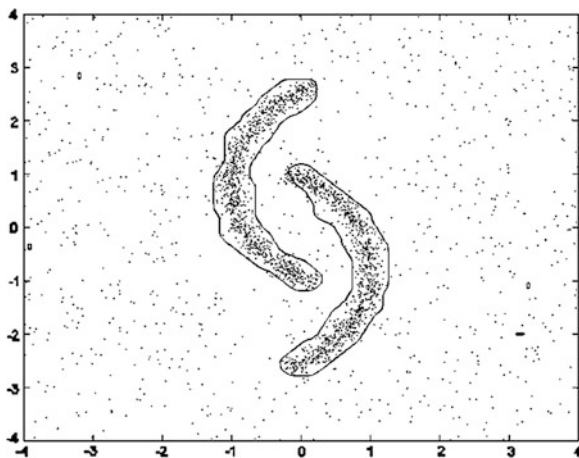
As shown in these examples, it can be seen that our approach has some advantages, for instant, our approach doesn't require the number of clusters and specific learning step, is suitable to discover clusters with arbitrary shape. Noises and isolated points are easy to be identified in the above results.

Since convolution is involved in Amari's neural field, which consumes a lot of computation, the computational time expense of this approach is high. As a result, limited by current computer technology, this approach cannot deal with high dimensional clustering problems efficiently. However, these Numeric examples show that this approach has high accuracy in clustering and anti-noise ability.

**Fig. 15.1** Three clusters obtained by this approach for three-Gaussian data set



**Fig. 15.2** Two non-convex clusters obtained by this approach for “Double C” data set with noises



Moreover, these clustering results are highly close to human cognition. So this approach shows potential especially when breakthroughs in computer technology like quantum computer are made in future.

## 15.5 Conclusion

In this paper, we present a new clustering approach inspired from human vision. By reproducing the mechanism of the three functional levels of human vision, we present a new clustering approach. This approach is biologically plausible, robust to noises and suitable to discover arbitrary shaped clusters. To show the feasibility of our approach, some numeric examples are given.

Nevertheless, our approach is an attempt. Our approach is suitable for all kinds of data sets theoretically, but it relies on a neural field which contains a convolution on its right hand side, which would consume much computer time.

**Acknowledgments** This paper is sponsored by the Scientific Research Foundation of Guangxi University (Grant No. XBZ120366) and supported by NSFC, Tian Yuan Special Foundation, Project No. 11226141.

## References

1. Xu R, Wunsch D (2005) Survey of clustering algorithms. *IEEE Trans Neural Netw* 16(3):645–678
2. Bishop CM (2006) *Pattern recognition and machine learning*. Springer, New York
3. Giese MA (1999) *Dynamic neural field theory for motion perception*. Kluwer Academic Publishers, Norwell

4. Kopecz K, Schoner G (1995) Saccadic motor planning by integrating visual information and pre-information on neural dynamic fields. *Biol Cybern* 73(1):49–60
5. Leung Y, Zhang J, Xu Z (2000) Clustering by scale-space filtering. *IEEE Trans Pattern Anal Mach Intell* 22(12):1396–1410
6. Xu Z, Meng D, Jing W (2005) A new approach for classification: visual simulation point of view. *Lecture notes in computer science advances in neural networks—proceedings of international symposium on neural networks, ISNN 2005, Part II, Chongqing*, pp 1–7
7. Amari S (1977) Dynamics of pattern formation in lateral-inhibition type neural field. *Biol Cybern* 27(2):77–87
8. Jin D, Peng J (2009) A new approach for estimating the attraction domain for Hopfield-type neural networks. *Neural Comput* 21(1):101–120