

# Image Segmentation with Multiple Hypergraph Fusion and Superpixels

Kaixiang Wang

School of Computer Science and Technology,  
Nanjing Normal University,  
Nanjing 210046, China

**Abstract.** - This paper presents a new method of image segmentation based on superpixels and multiple hypergraph fusion. In this paper, the original image is oversegmented by the ultra-pixel segmentation algorithm SLIC, and the super-pixel blocks are extracted from the two aspects of color information and gradient information. A hypergraph model is constructed for each channel of the features. From the respect of random walk, the information of multiple hypergraphs is fused to construct the multi-hypergraph laplacian matrix. After obtaining the multi-hypergraph laplacian matrix, we construct the optimization model and solve the model by crossing the iterations. Then we can get the results of the superpixel blocks and the image segmentation results. With the superpixels and multiple hypergraph fusion, our method can reduce the loss of information effectively and make the segmentation more precise. Our method was evaluated in Berkeley Segmentation Database. The comparative experiments result with some state-of-the-art segmentation algorithms show that our method can give better segmentations for the natural images.

**Keywords:** *Image Segmentation, Multi-Hypergraph, Superpixels, Spectral Clustering*

## 1. INTRODUCTION

With the development of information technology, image technology plays an increasingly important role in artificial intelligence, target recognition and tracking, medical imaging and other fields. Image technology consists of three parts, image processing, image analysis and image recognition, the image processing stage is the basis and key to the following applications. The image segmentation as an important step in the image processing stage is the focus of research on the whole image field. Therefore, image segmentation technology has become a very important and difficult research topic in pattern recognition, artificial intelligence and computer vision [25].

Image segmentation is the process of partitioning an image into objects and their backgrounds [26]. Image segmentation based on graph theory and hypergraph has always played an important role in image segmentation [3]. The theory of mathematic provides theoretical support for its development. The theory of graph provides a unified framework for any image. The structure of graph theory is very suitable for image segmentation, and the idea of image segmentation based on graph theory shows a good development prospect.

A graph-based representation can only represent binary relations between vertices, but we may need to represent a higher order relationship in many situations. An extension is provided by hypergraphs [27], which is a generalization of a graph, as the underlying model for the image content. The edges of hypergraphs called hyperedges, each hyperedge is a subset of the set of vertices. As an image modeling tool, compared to standard graphs that model pairs of pixels as their edges, hyper-graphs can model image patches as their hyperedges. By the means of hyperedges, higher-order relations between nodes can be directly modeled in a hypergraph [5].

The traditional structure of the graph-based segmentation algorithm is based on the pixels. With the size of image increasing, the scale of the graph increasing accordingly. If the superpixels are used instead of the pixels, the scale of the graph and the computational complexity can be greatly reduced. The superpixel is a set of pixels in the image that are consistent and capable of expressing the local structural features of images. The number of super pixels is much smaller than the number of pixels in the image, this will help to reduce the computational complexity of subsequent processing. The superpixel also has some ability of local structural feature expression, which is beneficial to the extraction and expression of local features.

Although segmentation algorithm based on graph theory has been widely applied, it still has some limitations. In this paper, we proposed a new strategy for image segmentation method based on superpixels and multiple hypergraph fusion, namely Multiple Hypergraph Fusion Segmentation(MHFS). Algorithm flowchart is show in Figure 1. Compared with the current graph-based segmentation algorithm, the proposed MHFS has the following contributions: Firstly, the normal graph can only depict the binary relation between objects, but the hypergraph structure we use can depict higher-order relationships; the hypergraph structure can express higher-order relationships between superpixels better and improve the accuracy and validity of the model. Secondly, we construct different hypergraphs based on different features, then we fused the hypergraphs through constructing a fusion laplacian matrix and fusing the hypergraphs in the view of random walk. This step reduces the loss of information effectively and makes the segmentation more precise. Thirdly, we use segmentation method based on superpixels for image segmentation, which effectively reduces the size of the graph model and the improves the efficiency of the algorithm.

The rest of this paper is organized as follows: The proposed image segmentation algorithm is introduced in section 2. In section 3, we evaluated our methods and compare with some state of art methods. Conclusions are discussed in section 4.

## 2. APPROACH

In this chapter, we will detail our proposed algorithm. We introduce some basic notions on hypergraphs before the main steps of our approach. The first step of our approach is to generate superpixels and extract the features. Our second step is to construct hypergraphs based on different features. Then we construct the Multi-hypergraph laplacian matrix. Finally, we conduct the multi-hypergraph spectral clustering to get the final result of segmentation.

### 2.1. Preliminary

Let  $V$  denote a finite set of samples, and let  $E$  be a family of subsets  $e$  of  $V$  such that  $\cup_{e \in E} e = V$ .  $G = (V, E)$  is then called a hypergraph with the vertex set  $V$  and the hyperedge set  $E$ . A hyperedge which contains just two vertices is just a simple graph edge. A weighted hypergraph is a hypergraph that has a positive number  $\omega(e)$  associated with each hyperedge  $e$ , called the weight of hyperedge  $e$ . We denote a weighted hypergraph by  $G = (V, E, \omega)$ . A hyperedge  $e$  is said to be incident with a vertex  $v$  when  $v \in e$ . The degree of each vertex  $v \in V$  is defined as:

$$d(v) = \sum_{\{e \in E | v \in e\}} \omega(e) \quad (1)$$

We let  $|S|$  denote the cardinality of a given arbitrary  $S$ . The degree of a hypergraph  $e \in E$  is defined to as:

$$\delta(e) = |e| \quad (2)$$

A hypergraph  $G$  can be represented by a  $|V| \times |E|$  matrix  $H$  with entries  $h(v, e) = 1$ , if  $v \in e$  and 0 otherwise, called the incidence matrix of  $G$ . Based on matrix  $H$ , the degree of each vertex and each hyperedge can be calculated as:

$$d(v) = \sum_{e \in E} \omega(e) h(v, e) \quad (3)$$

$$\delta(e) = \sum_{v \in V} h(v, e) \quad (4)$$

Let  $D_v$  and  $D_e$  denote the diagonal matrices containing the vertex and hyperedge degrees respectively, and let  $W$  denote the diagonal matrix containing the weights of hyperedges. Then the adjacency matrix  $A$  of hypergraph  $G$  is defined as:

$$A = HWH^T - D_v \quad (5)$$

where  $H^T$  is the transpose of  $H$ . [6]

### 2.2. Superpixels generation and feature extraction

In this paper, we adopt the simple linear iterative clustering (SLIC) [17] to over segmentation the image. We can get compact, nearly uniform ultra-pixel because SLIC has several advantages: 1) the generated super pixels as cells are usually compact and neat, from which neighborhood features are easier to express. 2) It can not only work on the color map, but also be compatible with grayscale. 3) It needs to set very few parameters, i.e., only the preset number of ultra-pixel. 4) compared to other ultra-pixel segmentation method, SLIC has advantages in the running speed, bringing ultra-pixel compactness and good contour retention.

When the superpixels are extracted, we can handle the image segmentation by the superpixels, which reduces the computing complexity greatly. In our methods, we employ two kinds of features commonly used in image segmentation: color and gradient information. Each of these two kinds of features has three channels. In our method, we will hypergraph for each channel, resulting in six hypergraphs. In the following, before we introduce how to construct hypergraphs, we show how to extract these two kind of features.

The RGB color space is widely used in image segmentation. Actually, it also has the simplest formulation and explanation. RGB color space to red, green and blue based on the three basic colors, to varying degrees of superposition, resulting in a rich and wide range of colors, so commonly known as three primary colors mode. The histogram of color often constructed on images, now we apply it to superpixels. For a color image, each superpixels has three channels. For each superpixel, we will generate three histograms according to three channels.

Besides the color features, gradient features can help to perform segmentation too. Histogram of oriented gradient based on superpixels is derived from the HOG characteristics, [24] we repute a superpixel here as a cell in the HOG, the superpixel block is irregular so we cannot get a block consist of several cells. Here we just use HOG to describe the cell direction gradient information to describe the super-pixel gradient information. For each super pixel block, we can get a vector representing the gradient information on each of the three channels.

Thus we have six vectors to describe the image features, and then we will create a super-map for each of the six features, respectively, from the perspective of the six features to describe the high-order relationship between the ultra-pixels.

### 2.3. Hypergraph construction

The INH (Image Neighborhood Hypergraph) model [5][8] has been found very effective at representing the image content. In this paper, we use the superpixels based INH to construct hypergraphs.

A hypergraph is a pair  $H = (X, E)$ , where  $X = X_1, X_2, \dots, X_N$  is the set of superpixels and  $E = E_1, E_2, \dots, E_N$  is the set of hyperedges. The image will be represented by the following mapping:

$$I : X \in Z^2 \rightarrow C \in Z^n(6)$$

Vertices of X are called superpixels, elements of C are called features. Let d be a distance on X, it defines the distance on space. Let d' be a distance on C, it describes the distance on feature.

Different from the original INH model, we conduct the model on superpixels and we use the mean of space location of pixels in a superpixel to represent the space location of the superpixel. The features of superpixels are constructed in section 3.1.

We have a neighborhood relation on an image defined by:

$$\forall x \in X, \Gamma_{\alpha, \beta}(x) = \{x' \in X, x' \neq x | d(I(X), I(X')) < \alpha \text{ and } d'(x, x') \leq \beta\} \quad (7)$$

The neighborhood of x on the grid will be denoted by  $\Gamma_{\alpha, \beta}(x)$ . To each image I, we can associate a hypergraph:

$$H_{\Gamma_{\alpha, \beta}}(I) = (X, (x \cup \Gamma_{\alpha, \beta}(x))) \quad (8)$$

#### 2.4. Multi-hypergraph laplacian matrix construction

From the above, we can build an image superpixel hypergraph structure for each feature. Then we develop a method to combine the information of the hypergraphs. Specifically, we associate each hypergraph with a natural random walk.[15] Then the transition probabilities and stationary distribution of the hypergraphs are fused.

Let P denote the transition probability matrix of this hypergraph random walk. Then each entry of P is:

$$p(u, v) = \sum_{e \in E} \omega(e) \frac{h(u, e) h(v, e)}{d(u) \delta(e)} \quad (9)$$

The stationary distribution  $\pi$  of the random walk is:

$$\pi(v) = \frac{d(v)}{\text{vol}(V)} \quad (10)$$

Then we can explain the multiple hypergraph cut in terms of a random walk[16]:

$$\beta_1(u) = \frac{\alpha \pi_1(u)}{\alpha \pi_1(u) + (1 - \alpha) \pi_2(u)} \quad (11)$$

$$\beta_2(u) = \frac{(1 - \alpha) \pi_2(u)}{\alpha \pi_1(u) + (1 - \alpha) \pi_2(u)} \quad (12)$$

The parameter  $\alpha$  is used to specify the relative importance of each graph in clustering. Then new transition probabilities among vertices are defined as:

$$p(u, v) = \beta_1(u) p_1(u, v) + \beta_2(u) p_2(u, v) \quad (13)$$

Then we can get the stationary distribution from a straightforward computation.

$$\pi(v) = \alpha \pi_1(v) + (1 - \alpha) \pi_2(v) \quad (14)$$

The above fusion is conduct on two hypergraphs, but we may have more than two hypergraphs to fuse in many time. In this paper, we have six hypergraphs to fuse for each image, so we need extend this algorithm and make it enable to fuse any numbers of hypergraphs.

We extend the number of graphs from two to N and transfer the equation to matrix notation:

$$P = \Pi^{-1} \sum_{i=1}^N \alpha_i \Pi_i P_i \quad (15)$$

$$\Pi = \sum_{i=1}^N \alpha_i \Pi_i \quad (16)$$

Please be noted that, as shown, the transition probability matrix and stationary distribution matrix of the fused hypergraph are not simply linear combinations on the laplacian matrix of each hypergraph.

Then we can get the laplacian matrix of the fused hypergraph as follows:

$$L = \Pi - \frac{\Pi P + P^T \Pi}{2} = \sum_{i=1}^N \alpha_i \Pi_i (I - P_i) \quad (17)$$

#### 2.5. Multi-hypergraph spectral clustering

Different machine learning tasks can be conducted on hypergraphs. The spectral clustering framework can be formulated as:

$$\min_{X \in R^{n \times k}} \text{Tr}(X^T D^{-1/2} L D^{-1/2} X) \quad (18)$$

s. t.  $X^T X = I$

where  $X \in R^{n \times k}$  is a matrix consisting of the column vectors and k is the number of clusters. The column vectors of X are indeed the eigenvectors corresponding to the k smallest eigenvalues obtained from the eigen-decomposition (EVD) on  $D^{-1/2} L D^{-1/2}$ . After row normalization of these eigenvectors, we can apply any classical clustering algorithm such as k-means to partition these low-dimensional embeddings.[12]

As mentioned in section 3.3, the multi-hypergraph laplacian matrix has been contracted. Then, the multiple hypergraph fusion based spectral clustering can be formulated as:

$$\begin{aligned} \min_{X \in R^{n \times k}, \alpha} \text{Tr} \left( X^T \left( \sum_{i=1}^N \alpha_i \Pi(I - P_i) \right) X \right) + \lambda \|\alpha\|^2 \quad (19) \\ \text{s. t. } \sum_{i=1}^N \alpha_i = 1, \alpha_i > 0, i = 1, \dots, N \\ X^T X = I \end{aligned}$$

The regularization term  $\|\alpha\|^2$  is adopted to avoid the fastening on one manifold, and  $\lambda$  is the trade-off parameter to control the contribution of regularization term  $\|\alpha\|^2$ . [9][13][14]

The optimization problem can be solved by cross iteration.

For a fixed  $\alpha$ , Eq.(19) can be simplified to:

$$\begin{aligned} \min_{X \in R^{n \times k}, \alpha} \text{Tr} \left( X^T \left( \sum_{i=1}^N \alpha_i \Pi(I - P_i) \right) X \right) + \lambda \|\alpha\|^2 \quad (20) \\ \text{s. t. } X^T X = I \end{aligned}$$

where  $M = \sum_{i=1}^N \alpha_i \Pi_i P_i$ . The column vectors of  $X$  are indeed the eigenvectors corresponding to the  $k$  smallest eigenvalues obtained from the eigen-decomposition (EVD) on  $M$ .

For a fixed  $X$ , Eq.(19) can be simplified to:

$$\begin{aligned} \min_{X \in R^{n \times k}, \alpha} -\text{Tr} \left( \sum_{i=1}^N \alpha_i M_i \right) + \lambda \|\alpha\|^2 \quad (21) \\ \text{s. t. } \sum_{i=1}^N \alpha_i = 1 \end{aligned}$$

Where  $M_i = X^T (\Pi_i P_i) X$ . The solution can be obtained as followed:

$$\alpha_i = \frac{N \text{Tr}(M_i) - \sum_{i=1}^N \text{Tr}(M_i) + 2\lambda}{2N\lambda} \quad (22)$$

We can get the  $X \in R^{n \times k}$  after the cross iterations above,  $X \in R^{n \times k}$  is a matrix consisting of the column vectors. After row normalization of these eigenvectors, we can apply any classical clustering algorithm such as  $k$ -means to partition these low-dimensional embeddings. Then the superpixels of the image will be divided to  $k$  parts.

### 3. EXPERIMENTS

In this section, the proposed MHFS (multiple hypergraph fusion segmentation) is extensively evaluated on the Berkeley Segmentation Database in comparison with the state of the art. We follow [10] to use the Berkeley Segmentation Dataset as the experimental data support. BSDS500 database includes 500 images and the corresponding ground truth data (each image has at least 4 human annotations), we choose 30 color images randomly from the BSDS500.

In this part of the experiment, we selected the Mean Shift[7], SAS[4], GL-graph[10] three image segmentation methods and our methods for comparative experiments. All results are compared to the human segmented segmentation results. In order to quantitatively compare the difference in performance of four algorithms, we introduce precision for regions, recall for regions, F-score for regions, IOU, error. Precision, recall, IOU, F-score values between 0 and 1, and the value gets closer to 1 means a better performance. Error also values between 0 and 1, and the values gets closer to 0 means a better performance.

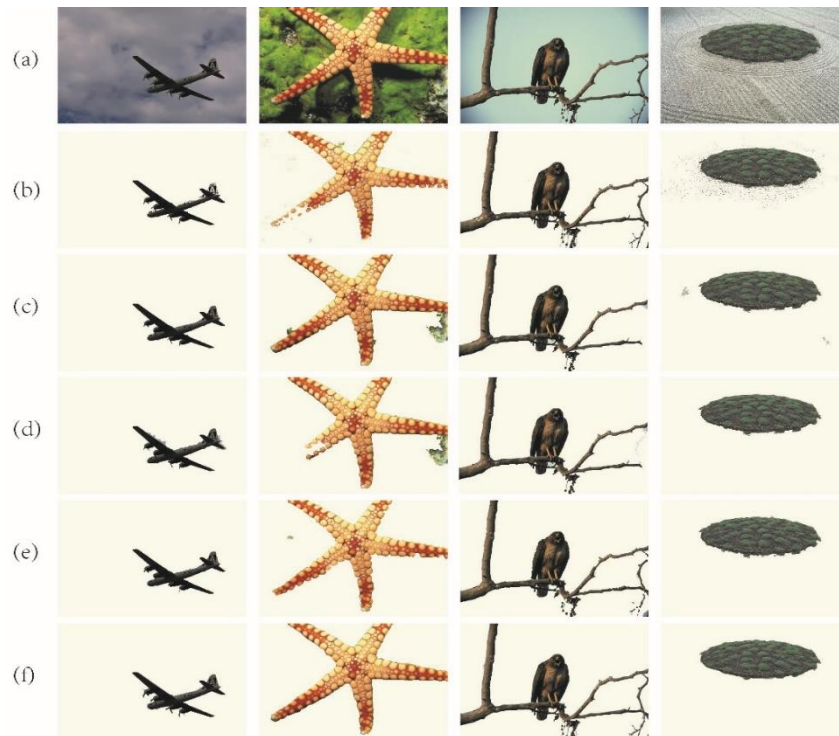


Fig. 2. Comparison of our method with state-of-art methods over the Berkeley segmentation database.

Figure 2 shows segmentation results of Meanshift, SAS, GL-graph and the proposed method. (a) shows the original image, (b) shows the results of MS, (c) shows the results of SAS, (d) shows the results of GL-graph, (e) is obtained with our method, and (f) is the groundtruth.

Table 1 records the results of segmentation precision for regions, recall for regions, F-score, IOU, error for the four algorithms. It is also the comparisons of our method(MHFS) with state-of-art methods over the Berkeley segmentation database.

Through the above numerical experiments, we can see that the proposed segmentation algorithm outperforms the Normalized Cut segmentation algorithm, the SAS segmentation algorithm, the GL-graph segmentation algorithm with performance and efficiency. From the experimental results, we can clearly find that the segmentation method we proposed can better deal with the edge of the image segmentation problem. Such as the first picture, there is a plane in the sky while the plane appearance similar color and the color of the sky, but we can segment them perfectly. The reason of performance is that the hypergraph structure can describe higher-order relationships, so it is very suitable to deal with edge problem. The GL-graph is based on the common graph model. The second order relationship of ordinary graph is difficult to depict the complex relationship between superpixels.

Table 1: Comparison of our method with state-of-art methods over the Berkeley segmentation database

Methods	Precision	Recall	F-score	IOU	error
MS	94.04	92.63	93.04	87.35	4.74
SAS	93.52	96.23	94.65	90.29	3.25
GL-graph	91.79	96.10	93.58	88.30	4.49
MHFS	96.60	94.97	95.37	91.40	2.68

Both in SAS and GL – graph, there is only a graph/hypergraph constructed to describe the relationships among pixels/superpixels. For different characteristics, the relationship between the pixel block is different, there must be loss of information if we only use a hypergraph to represent the image. With the help of multiple hypergraphs based on different characteristics, we can describe high order relationships well and effectively reduce the loss of information, so the segmentation results are better.

#### 4. CONCLUSION

This paper presents an image segmentation method based on superpixels and multiple hypergraph fusion. The method mainly improves the efficiency and precision of the image segmentation from the following aspects: We use segmentation method based on superpixels to reduces the size of the graph model and the improves the efficiency of the algorithm. The hypergraph structure which we used expresses the higher-order relationships can express the relationships between superpixels well. The most important is that we construct hypergraphs on each feature and fuse the hypergraphs through constructing a fusion laplacian matrix in the view of random walk. The successful fusion of different features reduces the loss of information effectively and makes the segmentation more precise.

The proposed method has some defects need to deal with: (1) the description of the superpixel features exist certain difficulties, some traditional image feature descriptors based on pixels are difficult to directly applied to superpixels. (2) the relationship between vertexes and hyperedges in hypergraph structure is binary (0 or 1), which belongs to hard segmentation; In some cases such as the images has blurred boundary, the soft segmentation based on probability is more suitable.

#### REFERENCES

- [1] Liu J, Wang C, Danilevsky M, et al. Large-scale spectral clustering on graphs[C]// International Joint Conference on Artificial Intelligence. AAAI Press, 2013:1486-1492.
- [2] A.Y. Ng, M.I. Jordan, and Y. Weiss. On spectral clustering: analysis and an algorithm. In *Advances in Neural Information Processing Systems 14*, Cambridge, MA, 2002. MIT Press
- [3] Shi J, Malik J. Normalized Cuts and Image Segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 22(8):888-905
- [4] Li Z, Xiao-Ming Wu, Chang S F. Segmentation using superpixels: A bipartite graph partitioning approach[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2012:789-796
- [5] Bretto A, Gillibert L. Hypergraph-Based image representation[C]// *Iapr International Conference on Graph-Based Representations in Pattern Recognition*. Springer-Verlag, 2005:1-11
- [6] Zhou D, Huang J. Learning with hypergraphs: clustering, classification, and embedding[C]// *International Conference on Neural Information Processing Systems*. MIT Press, 2006:1601-1608
- [7] Comaniciu D, Meer P. Mean Shift: A Robust Approach Toward Feature Space Analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5):603-619
- [8] Rital S, Cherifi H, Miguet S. Weighted Adaptive Neighborhood Hypergraph Partitioning for Image Segmentation[C]// *International Conference on Pattern Recognition and Image Analysis*. Springer-Verlag, 2005:522-531
- [9] Yu J, Cheng J, Wang J, et al. Transductive Cartoon Retrieval by Multiple Hypergraph Learning[M]// *Neural Information Processing*. Springer Berlin Heidelberg, 2012:269-276
- [10] Wang X, Tang Y, Masnou S, et al. A Global/Local Affinity Graph for Image Segmentation[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2015, 24(4):1399
- [11] Liu Q, Sun Y, Wang C, et al. Elastic Net Hypergraph Learning for Image Clustering and Semi-supervised Classification[J]. *IEEE Transactions on Image Processing*, 2016, PP(99):1-1
- [12] Wei W, Yang H, Wu Z, et al. A Novel Image Segmentation Algorithm Based on Multiple Features Fusion with Hypergraph and Superpixel[M]// *Foundations of Intelligent Systems*. Springer Berlin Heidelberg, 2014:709-718
- [13] Xu C, Tao D, Xu C. A Survey on Multi-view Learning[J]. *Computer Science*, 2013
- [14] Xia T, Tao D, Mei T, et al. Multiview Spectral Embedding[J]. *IEEE Transactions on Systems Man and Cybernetics Part B*, 2010, 40(6):1438-1446
- [15] Meila M, Shi J. A random walks view of spectral segmentation. *AISTATS*[J]. *Ai and Statistics*, 2001
- [16] Zhou D, Burges C J C. Spectral clustering and transductive learning with multiple views[C]// *Machine Learning, Proceedings of the Twenty-Fourth International Conference*. DBLP, 2007:1159-1166
- [17] Achanta R, Shaji A, Smith K, et al. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11):2274
- [18] Yu J, Tao D, Wang M. Adaptive hypergraph learning and its application in image classification.[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2012, 21(7):3262.
- [19] Karypis G, Aggarwal R, Kurnar V, et al. Multilevel hypergraph partition: applications in vlsi design[C]// *Design Automation Conference*. ACM, 1997:526-529.
- [20] Koyutürk M, Aykanat C. Iterative-improvement-based declustering heuristics for multi-disk databases ☆, ☆☆[J]. *Information Systems*, 2005, 30(1):47-70.
- [21] Ducournau A, Bretto A, Rital S, et al. A reductive approach to hypergraph clustering: An application to image segmentation[J]. *Pattern Recognition*, 2012, 45(7):2788-2803.
- [22] Ding L, Yilmaz A. Interactive image segmentation using probabilistic hypergraphs[J]. *Pattern Recognition*, 2010, 43(5):1863-1873.
- [23] Huang Y, Liu Q, Metaxas D. ]Video object segmentation by hypergraph cut[C]// *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. DBLP, 2009:1738-1745.
- [24] Dalal N, Triggs B. Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: *CVPR*[J]. 2005, 1(12):886-893.
- [25] Cheng H D, Jiang X H, Sun Y, et al. Color image segmentation: advances and prospects[J]. *Pattern Recognition*, 2001, 34(12):2259-2281.
- [26] Zhang Y J. *Advances in Image and Video Segmentation*[J]. *express shipping\**, 2006(3).
- [27] Agarwal S, Lim J, Zelnikmanor L, et al. Beyond Pairwise Clustering[J]. 2013, 2:838 - 845.