Contour Grouping by Clustering with Multi-feature Similarity Measure

Xue Bai, Siwei Luo, Qi Zou, and Yibiao Zhao

School of Computer and Information Technology
Beijing Jiaotong University
Beijing, China
bjtuxbai@gmail.com

Abstract. Contour grouping is a key task in computer vision domain. It extracts the meaningful objects information from low-level image features and provides the input for the further processing. There have been many techniques proposed over the decades. As a useful data analysis method in machine learning, clustering is a natural way for doing the grouping task. However, due to many complicated factors in natural images, such as noises and clutter in background, many clustering algorithms, which just use pairwise similarity measure, are not robust enough and always fail to generate grouping results that are consistent with the visual objects perceived by human. In this article, we present how the grouping performance is improved by utilizing multi-feature similarity under the information based clustering framework compared with other clustering methods using pairwise similarity.

1 Introduction

As an important task in computer vision, contour grouping takes the basic image features (e.g. edgels) as input and forms the object contours for further processing. This vision task can be seen as a clustering process with some predefined similarity measure if there is no prior information about the detected objects. According to the Gestalt Laws of perceptual organization [1], the similarity between any pair of edgels can be calculated, so most methods use pairwise similarity matrix as input for clustering procedure. However, in practice, due to the limitation of pairwise similarity on capturing global data structure, the unsupervised recognition process is very sensitive to the quality of feature description and is affected significantly by the noisy features in background. In [2], the multi-feature grouping cue is introduced. It can be defined over three or more data features and is considered to be more general and reliable.

The information-based clustering (Iclust) [3] provides more flexible descriptions on data relationships by utilizing collective rather than pairwise measures of similarity. For contour grouping, we propose to use the collective similarity measure, named multi-feature similarity, as the input information for clustering process, and we compare the grouping results with the one produced by pairwise

E.R. Hancock et al. (Eds.): SSPR & SPR 2010, LNCS 6218, pp. 415–422, 2010.

similarity. The experiments show that the grouping quality is obviously improved by using multi-feature similarity.

In the rest of this paper, the Iclust algorithm and its framework for multi-feature similarity measure is introduced in Sect. 2. In Sect. 3, we describe the contour grouping process based on Iclust with multi-feature similarity. The experimental results is presented in Sect. 4, and the last part is our conclusion.

2 Information Based Clustering

The information based clustering method formulate clustering as a tradeoff between maximizing the mean similarity of elements within a cluster and minimizing the complexity of the description provided by cluster membership. The goal of the algorithm is, for each data element i, finding a probabilistic assignment to clusters $P(C|i)$ that maximize the object function

$$\mathcal{F} = \langle s \rangle - TI(C; i) \ , \tag{1}$$

where $\langle s \rangle = \sum_{C=1}^{N_C} P(C)s(C)$ is the mean similarity of elements chosen at random out of each cluster, $I(C; i) = \frac{1}{N} \sum_{i=1}^{N} \sum_{C=1}^{N_C} P(C|i) \log \left[\frac{P(C|i)}{P(C)} \right]$ is the mutual information between the clusters variable C and elements variable i, and T is the Lagrange multiplier. Furthermore, $s(C)$ is defined as the average similarity among elements chosen out of a single cluster

$$s(C) = \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \cdots \sum_{i_r=1}^{N} P(i_1|C)P(i_2|C) \cdots P(i_r|C)s(i_1, i_2, \cdots, i_r) \ . \tag{2}$$

The similarity measure $s(i_1, i_2, \cdots, i_r)$ in above formulation is a collective measure of similarity among $r(r > 2)$ elements $i_1, i_2, \cdots, i_r$. It is very useful for contour grouping task when multi-feature grouping cues (e.g. cocircularity requires at least three data elements [2]) are involved. Thus, Iclust provides a good framework for describing more various data relations, not just limited to pairwise relation. We will demonstrate how the multi-feature similarity affects the grouping results through experiments in Sect. 4.

3 Contour Grouping

To do contour grouping, the first step is to obtain edges detected by an edge detector. Here, we use pb edge detector [4] to get the edge points and then use edgelink algorithm [5] to form small line segments (edgels). Figure 1 gives an example of edges and edgels of the same image. We now can take these edgels as basic features or elements for further grouping process.

(a) Original image (b) Binary edges (c) Edgels

Fig. 1. An example of edges and edgels generated from the same image. (a) original image; (b) edges detected by edge detector; (c) edgels formed by edgelink.

3.1 Multi-feature Grouping Cue

For each edgel, we compute the mean gray value of nearby area to represent that edgel. As an edgel reflects the change of gray value from one side to the other, we need to calculate the mean gray values of two areas on both sides of the edgel respectively. As shown in Fig. 2, the mid point of an edgel is firstly picked up, and we can obtain a certain size of square area (e.g. 6×6). So, taking the edgel as a borderline, the mean gray values on both sides of the edgel can be calculated by pixels within the area enclosed by the square and the borderline.

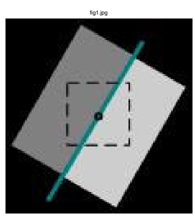


Fig. 2. Two regions beside an edgel

So far, we have two gray values for each edgel, one representing the area with small gray level and the other representing the area with higher gray level. To measure the similarity among multiple features or edgels, we first calculate the variance for each gray value among the edgels. It is inspired by the variance definition that "it is a measure of the dispersion of a sample". Then, the similarity value of $r (r > 2)$ edgels in terms of gray values is defined as following:

$$s(i_1, i_2, \cdots, i_r) = e^{-\frac{var_1^2(i_1,i_2,\cdots,i_r)}{\sigma_1^2}} \cdot e^{-\frac{var_2^2(i_1,i_2,\cdots,i_r)}{\sigma_2^2}}, \quad (3)$$

where $var_1(i_1, i_2, \cdots, i_r) = \frac{\sum_{i=1}^{r}(x_i - \overline{x})^2}{r-1}$, and $var_2(i_1, i_2, \cdots, i_r) = \frac{\sum_{i=1}^{r}(y_i - \overline{y})^2}{r-1}$. Variables x and y represent the two gray values respectively. The parameters σ_1

and σ_2 are the prior knowledge of the two variances. Here, we choose the average of all the values of var_1 for σ_1 and the average of all the values of var_2 for σ_2.

3.2 Clustering Process with Multi-feature Similarity Measure

As we mention in Sect. 2, the formulation of Iclust algorithm contains a collective measure of similarity which can describe the relation among multiple data elements. According to [3], if the derivative of object function (1) with respect to the variables $P(C|i)$ is equated to zero, we can obtain the optimal solution:

$$P(C|i) = \frac{P(C)}{Z(i;T)} \exp\left\{\frac{1}{T}[rs(C;i) - (r-1)s(C)]\right\} \ , \tag{4}$$

where $Z(i;T)$ is a normalization constant and $s(C;i)$ is the expected similarity between i and $r-1$ elements in cluster C. Equation (4) can be turned into an iterative algorithm that finds an explicit numerical solution for $P(C|i)$ corresponding to a (perhaps local) maximum of (1). Algorithm 1 gives the detailed procedure of the clustering algorithm. We implemented the algorithm based on [6], and for the calculation of $s(C)$ and $s(C;i)$, we only consider the "pure" multi-feature similarity $s(i_1, i_2, \cdots, i_r)$ in which the r elements $i_1, i_2, \cdots, i_r$ are different from each other.

Algorithm 1. Information-based clustering algorithm with multi-feature similarity

Input:
- parameter T and convergence parameter ϵ (we set $T = 1/25$ and $\epsilon = 1 \times 10^{-10}$ in our experiment)
- number of clusters N_C
- number of elements r in similarity measure $s(i_1, i_2, \cdots, i_r)$

Output: "soft" partition of the N elements into N_C clusters.

Initialization:
1. m=0
2. For each element $i = 1, \cdots, N : P^{(m)}(C|i) \leftarrow$ random distribution.

While True

For each element $i = 1, \cdots, N$:
1. update $P^{(m)}(C|i)$:
   $P^{(m+1)}(C|i) = \frac{P^{(m)}(C)}{Z(i;T)} \cdot \exp\left\{\frac{1}{T}[rs^{(m)}(C;i) - (r-1)s^{(m)}(C)]\right\}$
   where
   - $s(C;i) = \sum_{i_1=1}^{N} \cdots \sum_{i_{r-1}=1}^{N} P(i_1|C) \cdots P(i_{r-1}|C)s(i_1, \cdots, i_{r-1}, i)$
   - $s(C) = \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \cdots \sum_{i_r=1}^{N} P(i_1|C)P(i_2|C) \cdots P(i_r|C)s(i_1, i_2, \cdots, i_r)$
   - $P(i|C) = P(C|i)P(i)/P(C)$, $P(C) = \sum_{i=1}^{N} P(C|i)P(i)$, $P(i) = 1/N$
   - $Z(i;T) = \sum_{C'=1}^{N_C} P(C'|i)$
2. m=m+1
3. if $\left|P^{(m+1)}(C|i) - P^{(m)}(C|i)\right| \leq \epsilon$, break.

4 Experimentation and Analysis

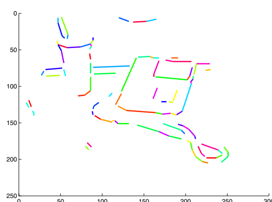
In our experiment, we pick up gray images and only consider gray value information when calculate similarities described in Sect. 3.1. To investigate the influence of multi-feature similarity measure on grouping performance, we test two clustering algorithms using pairwise similarity. One is Ncut [7], which is considered an effective clustering method for perceptual organization of low-level image features by partitioning a graph representation [8]. The other is also based on Iclust in the case of parameter $r = 2$. We define the pairwise similarity measure based on mean gray values of areas beside an edgel as well:

$$s(i_1, i_2) = e^{-\frac{d_1^2(i_1,i_2)}{\sigma_1^2}} \cdot e^{-\frac{d_2^2(i_1,i_2)}{\sigma_2^2}} , \qquad (5)$$

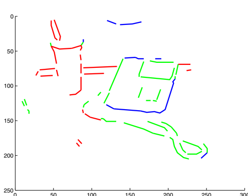
where $d_1(i_1, i_2) = |x_1 - x_2|$, and $d_2(i_1, i_2) = |y_1 - y_2|$. Variables x and y represent the two gray values respectively. The parameters σ_1 and σ_2 are the prior knowledge of two distance measures d_1 and d_2. Specifically, we choose the average of all the values of d_1 for σ_1 and the average of all the values of d_2 for σ_2. For grouping with multi-feature similarity measure, we test the case where $r = 3$. As mentioned in Sect. 3.2, only the "pure" 3-feature similarities $s(i_1, i_2, i_3)$ (i_1, i_2, i_3 are different edgel elements) are considered. Thus, the clustering process is totally based on multi-feature similarity.

We first test whether the clustering procedures can distinguish different object contours in the image including two salient objects. Figure 3 shows that Ncut and Iclust with 3-feature similarity have better performance than Iclust with pairwise similarity.

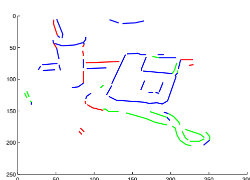
(a) Original image



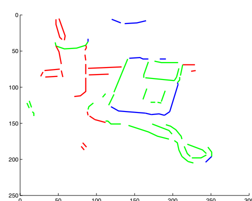
(b) Detected edgels



(c) Ncut

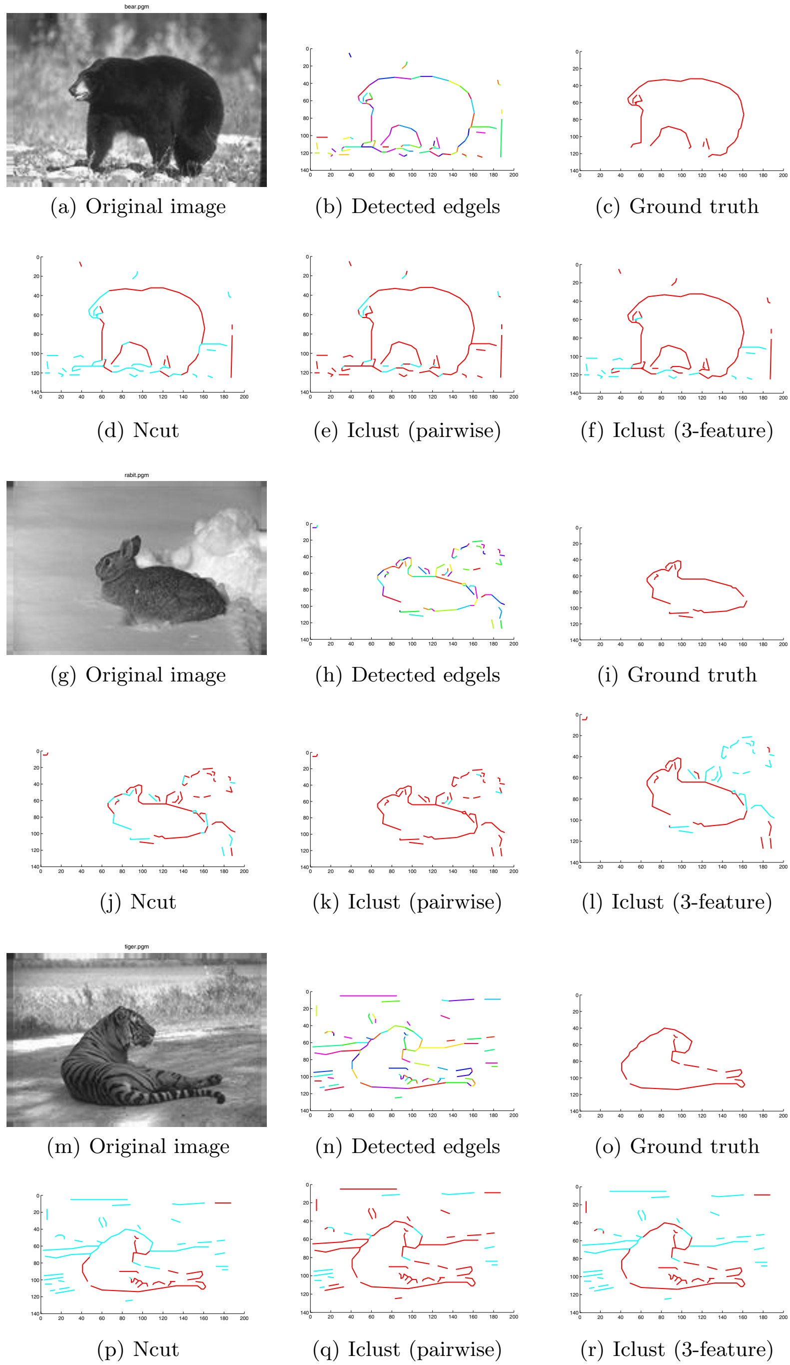


(d) Iclust (pairwise)



(e) Iclust (3-feature)

Fig. 3. Grouping results for the image with multiple objects



(a) Original image (b) Detected edgels (c) Ground truth

(d) Ncut (e) Iclust (pairwise) (f) Iclust (3-feature)

(g) Original image (h) Detected edgels (i) Ground truth

(j) Ncut (k) Iclust (pairwise) (l) Iclust (3-feature)

(m) Original image (n) Detected edgels (o) Ground truth

(p) Ncut (q) Iclust (pairwise) (r) Iclust (3-feature)

Fig. 4. Grouping results produced by Ncut, information-based clustering with pairwise similarity and information-based clustering with 3-feature similarity

Table 1. Grouping performance measure

Image Label	Performance Measure	Ncut	Iclust(pairwise)	Iclust(3-feature)
Fig.4 (a)	Precision	0.78	0.50	0.60
	Recall	0.54	0.85	0.89
	β	0.65	0.65	**0.73**
Fig.4 (g)	Precision	0.41	0.47	0.78
	Recall	0.70	1.00	0.95
	β	0.54	0.68	**0.86**
Fig.4 (m)	Precision	0.57	0.44	0.59
	Recall	0.64	0.92	0.85
	β	0.60	0.64	**0.71**

Figure 4 shows the grouping results for extracting one salient object contour from background. We evaluate the grouping quality by calculating precision and recall values, and the total performance is measured by $\beta = \sqrt{precision \cdot recall}$. Table 1 lists the performance measure for each clustering procedure. As we just consider gray information in calculating similarities, and a gray value is an average gray level within a certain small area, this feature description is relatively rough and could bring some inaccuracy to the measurement of similarity. We observe that 3-feature similarity measure has more stable and better grouping performance than the other two clustering procedures using pairwise similarity. It indicates that multi-feature similarity is more robust and not sensitive to the quality of feature description.

5 Conclusion

We present how the multi-feature similarity measure influence the grouping results under the information-based clustering framework for the computer vision task of contour grouping. We define this kind of similarity based on the variance of gray values over multiple edgels. Through the experiment, we find that multi-feature grouping cue is more reliable and robust compared with bi-feature cue. In the future work, more image data and various values of parameter r should be tested. And beside using variance for calculating multi-feature similarity, other grouping cues should be further investigated.

Acknowledgments

This work is supported by National High Technology Research and Development Program of China (2007AA01Z168), National Nature Science Foundation of China (60975078, 60902058, 60805041, 60872082, 60773016), Beijing Natural Science Foundation (4092033) and Doctoral Foundations of Ministry of Education of China (200800041049).

References

1. Elder, J.H., Goldberg, R.M.: Ecological Statistics for the Gestalt Laws of Perceptual Organization of Contours. Journal of Vision 2(4), 324–353 (2002)
2. Amir, A., Lindenbaum, M.: A Generic Grouping Algorithm and Its Quantitative Analysis. IEEE Trans. Pattern Analysis and Machine Intelligence 20(2), 168–185 (1998)
3. Slonim, N., Atwal, G.S., Tkačik, G., Bialek, W.: Information-based clustering. Proceedings of the National Academy of Sciences of the United States of America 102(51), 18297–18302 (2005)
4. Martin, D.R., Fowlkes, C.C., Malik, J.: Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(5), 530–549 (2004)
5. Kovesi, P.D.: MATLAB and Octave Functions for Computer Vision and Image Processing. School of Computer Science & Software Engineering, The University of Western Australia (2000), `http://www.csse.uwa.edu.au/~pk/research/matlabfns/`
6. Slonim, N., Atwal, G.S., Tkačik, G., Bialek, W.: Iclust_1.0: Matlab Code for Information Based Clustering (2004), `http://www.princeton.edu/~nslonim`
7. Shi, J., Malik, J.: Normalized Cuts and Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8), 888–905 (2000)
8. Soundararajan, P., Sarkar, S.: An In-Depth Study of Graph Partitioning Measures for Perceptual Organization. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(6), 642–660 (2003)