Analysis of MinCut, Average Cut, and Normalized Cut Measures (Extended Abstract)

Padmanabhan Soundararajan and Sudeep Sarkar[1]
Computer Science and Engineering
University of South Florida, Tampa, FL 33620
Email: {psoundar, sarkar}@csee.usf.edu

Partitioning of a graph representation, defined over low-level image features based on Gestalt inspired relations, is an effective strategy for forming coherent perceptual groups in an image. The usual practice, mainly motivated by efficiency considerations, is to approximate the general K-way partitioning solution by recursive bi-partitioning, where at each step the graph is broken into two parts based on a partitioning measure. We concentrate on four such measures, namely, the minimum cut [9], average cut [4], Shi-Malik normalized cut [5], and a variation of the Shi-Malik normalized cut. The minimum cut partition seeks to minimize the total link weights cut. The average cut measure is proportional to the total link weight cut, normalized by the sizes of the partitions. The Shi-Malik normalized cut measure is also a normalized measure, but the normalizing factor is the product of the total connectivity (valency) of the nodes in each partition. A natural variation of the Shi-Malik normalized cut measure, which one might suggest, is the total edge weight cut normalized by the total association in each partition. The questions we ask in this work are: Does the nature of the cut measure really matter? Are the quality of the groups significantly different for each cut measure? How do the measures vary in the space of all possible partitions? Other studies that also consider similar questions are those by Perona and Freeman [3], Amir and Lindenbaum [1], Weiss [7], who studied similarities of graph spectral methods for segmentation and Williams and Thornber [8]. Our study complements and extends the previous studies in that it (i) considers bi-partitioning in the presence of K-objects in the scene instead of just two, (ii) relates the measures to the underlying image statistics in a probabilistic manner, and (iii) undertakes a rigorous and extensive empirical evaluation.

Theoretical Analyses: In the first part of the study we analytically relate the nature of each of the partitioning measures to the underlying image statistics. This lets us quantify under what image statistics each measure would give us the correct partition. We denote the probability of association between features from the i-th object by p_{ii} and the probability of association between features from *different* objects by p. Let the number of features of the i-th object be denoted by N_i and the total number of image features be N. *We assume that the probability of association of inter-object features is lower than intra-object features, i.e $p < \min(p_{11}, \cdots, p_{KK})$, where K is the number of objects (background is considered as one object).* Let a bi-partition result in two partitions such that $f_i N_i$ features from each object are in one partition (S_1) and the rest are in the other partition (S_2). Thus, any bi-partitioning cut is characterized by the column vector, $\mathbf{f} = (f_1, \cdots, f_K)^T$. For recursive bi-partitioning to result in the correct cut, f_i should be always 0 or 1, excluding the case when all f_i's are 0 or 1.

[1]This work was supported by the National Science Foundation grant Nos. IIS-9907141 and EIA-9729904.

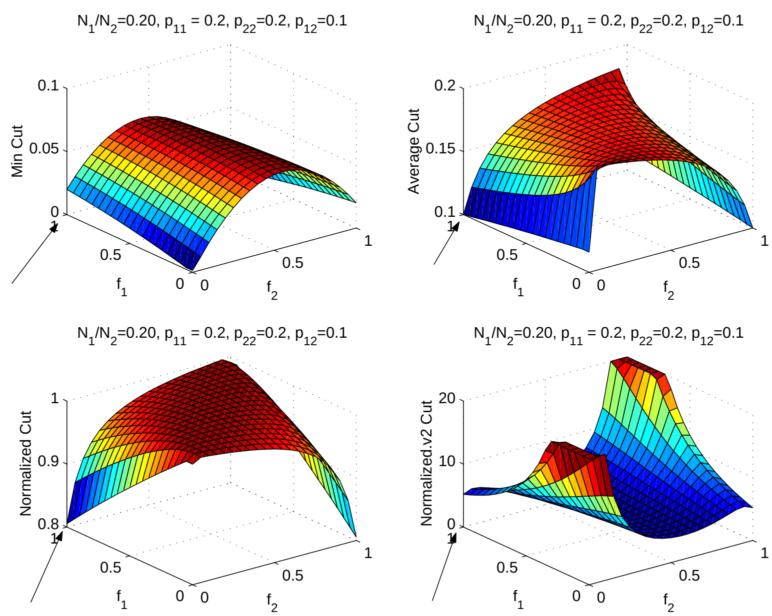

Figure 1. The expected values of the four measures, (a) minimum cut, (b) average cut, (c) Shi-Malik normalized cut, and (d) a variation of the Shi-Malik normalized cut measure plotted as a function of f_1 and f_2 for a scene with two *dissimilar sized objects* **and with the strength of connection within objects being just** *twice* **the strength between objects. The desired partition, namely $(f_1 = 1, f_2 = 0)$, is represented by the corner of the space indicated by an arrow. Notice that both the average cut and the normalized cut measures seem to be well formed with a minimum at the right partition.**

Analytical Conclusions

- As one might expect, the use of the minimum cut measure will not always result in correct partitioning. It will result in the correct solution only if the probability of connections between inter-object features is smaller than the ratio of the probability of connections intra-object features to the total number of features, excluding the object under consideration. The required condition is $p < \min\{\frac{p_{11}}{N-N_1}, \cdots, \frac{p_{KK}}{N-N_k}\}$, where N is the total number of image features.
- The minimization of the average cut and the normalized cut measure will result in the correct segmentation.
- While the average cut measure is not necessarily a concave function over the domain of possible partitions, the min-cut and the Shi-Malik normalized cut measure seems to be a concave function; this is a conjecture at this point.
- Recursive bi-partitioning using the average and the (Shi-Malik) normalized cut measure will be successful in segmenting K objects.

- A variation of the normalized cut measure, where we normalize by the association within each group, is not a well formed measure. It will result in wrong groups for certain image statistics. In fact, it can be shown that there is a local extremum of this cut measure at $(f_1 = 0.5, \cdots, f_K = 0.5)$ with a value of 4, irrespective of the image statistics. And, this extremum value, i. e. 4, is less than the value of the measure at the correct partitions for $\frac{p_{ii} N_i}{p(N - N_i)} < 0.5$. This condition would be true, for instance, when we have two objects with different number of features, e.g. $N_1 < N_2$, and association between features from different objects, p, although weak is not sufficiently weak as compared to associations between features of the smaller object.

Empirical Evaluation: In the second part of the work, we empirically evaluate the groups produced by graph partitioning for three measures, viz. mincut, average cut, and Shi-Malik normalized cut, given the task of grouping constant curvature edge segments. The grouping framework is same as that described in our earlier work [4]. The nodes of the graph represent constant curvature segments and the links are quantified by the Gestalt inspired relations of parallelism, continuity, T-junction, L-junction, region similarity, and proximity. Link quantification is achieved using Bayesian networks [4]. The probabilities of the Bayesian networks include the priors and the conditional probabilities, which are parameterized in terms of threshold parameters. These probabilities are learnt from examples.

The recursive bi-partitionings of the graph based on the normalized cut and the average cut measures are approximated by the second eigenvector of the normalized Laplacian matrix and plain Laplacian matrix representation of the graph, respectively. We use the LEDA-4.0 implementation to compute the minimum cut [6].

Performance Measure: We evaluate the quality of the final groups in terms of the expected performance of a constrained search based object recognition system built based on the groups. In [2], Borra and Sarkar proposed performance measures can be computed *without* building a full vision system but instead, use manually specified ground truth. Let N_{G_i} denote the number of features in G_i the detected group, N_{O_j} denote the number of model features (ground truth) in O_j, and $N_{G_i \cap O_j}$ denote the number of group features that lie on the model. Two performance measures that capture the expected time and quality of model matching are: $P_{time}(G_i, O_j) = \frac{N_{G_i \cap O_j}}{N_{G_i}}$ and $P_{qual}(G_i, O_j) = \frac{N_{G_i \cap O_j}}{N_{O_j}}$. Let $N_{overlaps}$ be the total number of overlaps between the computed groups and objects in a scene. We form a scalar performance measure, β, by combining these measure such that (i) it penalize a group that is spread across two objects more than a group that overlaps with one object and the background, (ii) large groups are preferred to small groups, and (iii) it penalizes groups of features that do not belong to any object.

Image Set: We use a database of 50 images. There are 10 images of natural objects in indoor backgrounds, 10 images of natural objects in outdoor backgrounds, 10 images of man-made objects in outdoor background, 10 images of man-made objects in indoor backgrounds, and 10 aerial images that contained man-made objects in natural surroundings. So essentially, we have 5 different image classes. Each of these images is associated with manual outlines of objects of interest.

Parameter Selection: One of the important aspects of empirical performance evaluation is the choice of the parameters of the algorithm. The present grouping framework has 21 parameters: 7 are used in the Bayesian network to quantify the graph links, 6 prior probabilities for the relations, 2 parameters for the graph spectral partitioning (minimum cluster size and maximum partition strength), 3 edge detection parameters, 3 parameters are used for the constant curvature contour segmentation algorithm. We select the optimal parameters by stochastically sampling the parameter space using a team of learning automata, as described in [4]. From the stochastic sampling, we select 100 best performances on each of the 50 images for each of the 3 cut types.

Insights Derived:

1. The best possible performances on each image for the three different cut measures are not significantly different from each other. Analysis of Variance (ANOVA) studies show that although the variation of performance due to the cut type is not significant, the variation due to the image classes is significant. The interaction between the image class and the cut type is also not significant. Non parametric tests such as Wilcoxon Rank Sum Scores also lead to the conclusion that the best performances of each cut type are statistically similar.
2. The performance with the three measures is worst for man made objects in indoor surroundings and for aerial images.
3. The range of performance for the best 100 parameters for the three cuts are not statistically different.
4. The mean performance with the 100 best parameters for the three cut measures are also not statistically significant.
5. The normalized cut based partition took significantly larger time to compute than the average and the mincut.

References

[1] A. Amir and M. Lindenbaum. A generic grouping algorithm and its quantitative analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:168–185, Feb 1998.

[2] S. Borra and S. Sarkar. A framework for performance characterization of intermediate-level grouping modules. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:1306–1312, Nov 1997.

[3] P. Perona and W. Freeman. A factorization approach to grouping. *European Conference on Computer Vision*, pages 655–670, 1998.

[4] S. Sarkar and P. Soundararajan. Supervised learning of large perceptual organization: Graph spectral partitioning and learning automata. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(5):504–525, May 2000.

[5] J. Shi and J. Malik. Normalized cuts and image segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 731–737, Jun 1997.

[6] M. Stoer and F. Wagner. A simple min cut algorithm. *Algorithms - ESA '94*, pages 141–147, 1994.

[7] Y. Weiss. Segmentation using eigenvectors: A unifying view. In *Proceedings of the International Conference on Computer Vision*, volume 2, pages 975–982, 1999.

[8] L. R. Williams and K. K. Thornber. A comparison of measures for detecting natural shapes in cluttered backgrounds. *International Journal of Computer Vision*, 34(2/3):81–96, August 1999.

[9] Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: theory and application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:1101–1113, Nov 1993.