Investigation of Measures for Grouping by Graph Partitioning *

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

Abstract

Grouping by graph partitioning is an effective engine for perceptual organization. This graph partitioning process, mainly motivated by computational efficiency considerations, is usually implemented as recursive bi-partitioning, where at each step the graph is broken into two parts based on a partitioning measure. We study four such measures, namely, the minimum cut [11], average cut [6], Shi-Malik normalized cut [7], and a variation of the Shi-Malik normalized cut. Using probabilistic analysis we show that the minimization of the average cut and the normalized cut measure, using recursive bi-partitioning will, on an average, result in the correct segmentation. The minimum cut and the variation of the normalized cut will, on an average, not result in the correct segmentation and we can precisely express the conditions. Based on a rigorous empirical evaluation, we also show that, in practice, the quality of the groups generated using minimum, average or normalized cuts are statistically equivalent for object recognition, i.e. the best, the mean, and the variation of the qualities are statistically equivalent. We also find that for certain image classes, such as aerial and scenes with man-made objects in man-made surroundings, the performance of grouping by partitioning is the worst, irrespective of the cut measure.

1. Introduction

Partitioning of a graph representation, defined over low-level image features and 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 [11], average cut [6], Shi-Malik normalized cut [7], and a variation of the Shi-Malik normalized cut. The minimum cut partition seeks to minimize the total link weight cut. The average cut measure is proportional to the total link weight cut, normalized by the sizes of the partitions. The Shi-Malik 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 product of the total associations in each partition. The questions we ask in this work are: Do the nature of the cut measures 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? Are there classes of images on which grouping by partitioning does not work well?

The contributions of this paper are three fold. First, we analytically relate the nature of each of the partitioning measures to the underlying image statistics. This lets us quantify under what conditions each measure would give us the correct partitions. We find that the average cut measure and the Shi-Malik normalized cut measure will, on an average, result in the correct partition, whereas the minimum cut and the variation of the normalized cut do not. Second, we derive conditions under which the recursive bi-partitioning strategy, based on each of partitioning measures, results in correct partitions for scenes with more than two objects. We find that recursive partitioning with both the average cut and the normalized cut measures will, on an average, result in the correct partitions. Third, we empirically evaluate the groups produced by graph partitioning based on the three measures, viz. min-cut, average cut, and normalized cut, given the task of grouping extended edge segments. Our findings in this regards suggest that the quality of the groups with each of these three measures are statistically equivalent, as far as object recognition is concerned. We also examine whether the performance of the grouping-by-partitioning strategy depend on the image class.

Three studies that also considered comparison of different graph clustering methods are those of Weiss [9], who studied similarities of graph spectral methods for segmentation, and Williams and Thornber [10], who considered clustering methods based on the affinity matrix and Matula [3], who considered clustering methods based on the proximity

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

matrix. Our study complements and extends the previous studies in that we (i) consider K-objects in the scene, instead of just two, for the theoretical analysis, (ii) relate the cut measures to the underlying image statistics in a probabilistic manner, and (iii) undertake a rigorous and extensive empirical evaluation. The empirical evaluation is conducted on real images, using a rigorous automated parameter selection process, and thorough statistical analyses.

2. Theoretical Analysis

We consider for what kinds of image statistics would a measure, on an average, result in correct partitioning? In our analysis, we assume that we have an unweighted scene structure graph, as opposed to a weighted graph. This helps keep the mathematics tractable without compromising the essential insights into the partitioning process.

Definition 1 *Let*

1. *the number of objects be denoted by K: $(O_1, \cdots, O_K)$,*
2. *the number of features of the i-th object be N_i,*
3. *probability of existence of a link between features from the i-th object be p_{ii}, and*
4. *the probability of existence of a link between features from different objects be p.*

We can assume that the probability of association between inter-object features is lower than that between intra-object features. Thus, we have

Lemma 1 $p < \min(p_{11}, \cdots, p_{KK})$

Definition 2 *Let a bi-partition result in two partitions (S_1 and S_2) such that $f_i N_i$ features from the i-th object are in one partition (S_1) and the rest of the $(1 - f_i)N_i$ features are in the other partition (S_2). A bi-partitioning cut is thus characterized by the column vector: $\mathbf{f} = (f_1, \cdots, f_K)^T$.*
Note that f_is are discrete numbers that range from 0 to 1 in increments of $\frac{1}{N_i}$. For recursive bi-partitioning to result in the correct K-way cut, f_i should be always 0 or 1, excluding the case when all f_i's are 0 or all f_i's are 1.

It is trivial to show that,

Corollary 1 *The partition $(f_1, \cdots, f_K)$ is equivalent to the partition $(1 - f_1, \cdots, 1 - f_K)$.*

Definition 3 *Let $\mathbf{0}$ and $\mathbf{1}$ denote vectors whose components are all 0 and 1 respectively.*

Definition 4 *Let S denote the set of vectors $\mathbf{s}$, each of whose components, s_i is either 0 or 1, excluding the vectors $\mathbf{0}$ and $\mathbf{1}$. The dimension of $\mathbf{s}$ is the same as that of $\mathbf{f}$.*

Definition 5 *Let Φ denote the set of vectors ϕ, whose i-th components is either 0 or $\frac{1}{N_i}$, excluding the vector $\mathbf{0}$.*

Definition 6 *Let Ψ denote the set of vectors ψ, whose i-th components is either 1 or $1 - \frac{1}{N_i}$, excluding the vector $\mathbf{1}$.*

Definition 7 *The corner points on the boundary of the domain of possible partitioning cuts is given by the set $\{S \cup \Phi \cup \Psi\}$.*

That the elements of S will be boundary corner points is obvious. This set, S represents the desired partitions. The elements in the sets Ψ and Φ arise because we exclude $\mathbf{f} = \mathbf{0}$ and $\mathbf{f} = \mathbf{1}$, which do not represent a "partition", from the domain. We also make use of the fact that the possible values for f_i are $\frac{k}{N_i}$ for $k = 0, 1, \cdots, N_i$. Partitions represented by the elements in Ψ and Φ are undesirable partitions that separate just one feature of some object(s) from the others.

We next present the expressions of the expected number of links cut, the expected sizes of the partitions, the association within each partition, and the connectivity of the features in each partition using the above notations [1].

Theorem 1 *The expected number of links cut by a partitioning cut $\mathbf{f}$ is given by $Cut_{S_1,S_2}(\mathbf{f}) = \mathbf{f}^T\mathbf{P}(\mathbf{1} - \mathbf{f})$, where $\mathbf{P}$ is a $K \times K$ matrix with*

$$\mathbf{P}(i,j) = \begin{cases} p_{ii}N_i^2 & \text{for } i = j \\ pN_iN_j & \text{for } i \neq j \end{cases} \quad (1)$$

Theorem 2 *The number of features in each partition can be expressed as $Size_{S_1}(\mathbf{f}) = \mathbf{f}^T\mathbf{N}$ and $Size_{S_2}(\mathbf{f}) = (\mathbf{1} - \mathbf{f})^T\mathbf{N}$, where $\mathbf{N}$ is a column vector whose i-th entry is N_i,*

Theorem 3 *The expected numbers of total links within each partition are given by $Assoc_{S_1}(\mathbf{f}) = 0.5\mathbf{f}^T\mathbf{P}\mathbf{f}$ and $Assoc_{S_2}(\mathbf{f}) = 0.5(\mathbf{1} - \mathbf{f})^T\mathbf{P}(\mathbf{1} - \mathbf{f})$ where $\mathbf{P}$, is as in Theorem 1*

Theorem 4 *The expected connectivity of features in each partition, or the association of each partition with respect to the whole graph can be expressed as, $\mathrm{Connec}_{S_1}(\mathbf{f}) = \mathbf{1}^T\mathbf{P}\mathbf{f}$ and $\mathrm{Connec}_{S_2}(\mathbf{f}) = \mathbf{1}^T\mathbf{P}(\mathbf{1} - \mathbf{f})$ where $\mathbf{P}$ is as in Theorem 1*

Using the above expressions, we can express the four partitioning measures as follows

1. The minimum cut strategy simply tries to minimize the number of links cut by a partition. Thus, its expected value is

$$\mathrm{MinCut}_{S_1,S_2}(\mathbf{f}) = \mathbf{f}^T\mathbf{P}(\mathbf{1} - \mathbf{f}) \quad (2)$$

2. The average cut measure of a partition can be expressed as

$$\mathrm{AvgCut}_{S_1,S_2}(\mathbf{f}) = \mathbf{f}^T\mathbf{P}(\mathbf{1} - \mathbf{f})\left(\frac{1}{\mathbf{f}^T\mathbf{N}} + \frac{1}{(\mathbf{1} - \mathbf{f})^T\mathbf{N}}\right) \quad (3)$$

[1] Please note that due to space limitations some proofs are omitted in this section.

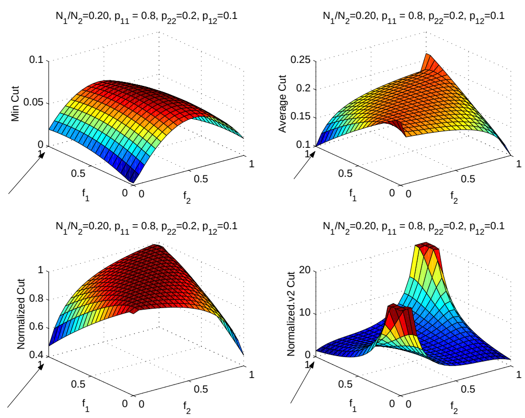


Figure 1. The expected values of the four measures, (a) minimum cut, (b) average cut, (c) normalized cut, and (d) a variation of the normalized cut measure plotted as a function of f_1 and f_2 for a scene with *dissimilar sized objects* and with the strength of connection within the smaller object being *eight* time the strength between objects and that for the larger object being just *twice*. $N_1 = 0.2N_2, p_{11} = 0.8, p_{22} = 0.2$, and $p = 0.1$.

3. The (Shi and Malik) expected value of the normalized cut cost can be expressed as follows:
$$\text{NormCut}_{S_1,S_2}(\mathbf{f}) = \mathbf{f}^T\mathbf{P}(\mathbf{1}-\mathbf{f})\left(\frac{1}{\mathbf{1}^T\mathbf{P}\mathbf{f}} + \frac{1}{\mathbf{1}^T\mathbf{P}(\mathbf{1}-\mathbf{f})}\right) \quad (4)$$

4. The expected value of the variation of the normalized cut cost can be expressed as:
$$\text{NormCutV2}_{S_1,S_2}(\mathbf{f}) = \begin{array}{l}\mathbf{f}^T\mathbf{P}(\mathbf{1}-\mathbf{f}) \\ \times\left(\frac{2}{\mathbf{f}^T\mathbf{P}\mathbf{f}} + \frac{2}{(\mathbf{1}-\mathbf{f})^T\mathbf{P}(\mathbf{1}-\mathbf{f})}\right)\end{array} \quad (5)$$

Fig. 1 shows the nature of the variation of the four cut measures as a function of all possible partitions, (f_1, f_2), of an image with two object, for one possible image statistics. In each figure, 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. While the minimum cut and the variation of the normalized cut does not always have a minimum at the correct partition. In the next subsections, we analytically establish these observations on a firmer footing.

2.1. Minimum Cut

Using the next two theorems, we establish the conditions under which the min-cut formalism is expected to result in the correct partitions by recursive bi-partitioning.

Theorem 5 *The expected number of links cut, $Cut_{S_1,S_2}(\mathbf{f})$, is a concave function.*

Proof: Omitted

Theorem 6 *The minimum cost partition, represented by $\mathbf{f_m}$, would maximize $|Size_{S_1}(\mathbf{f_m}) - Size_{S_2}(\mathbf{f_m})|$.*

Proof: We prove this by contradiction. Let us assume that $\mathbf{f}$ is a minimum cost partition whose corresponding $|\text{Size}_{S_1}(\mathbf{f_m}) - \text{Size}_{S_2}(\mathbf{f_m})|$ is not the maximum possible. We show that we can produce a lower cost partition with larger difference than this assumed minimum partition.

Let us consider the k-th component of $\mathbf{f}$, representing the partition of the features from the k-th object. The total partition cut cost can be expressed as the sum of two types of terms: the partition cost terms that include f_k and those that do not. We denote the aggregate of the cost terms that do not include k-th object by K. The total expected cost is:

$$\begin{aligned}\text{Cut}_{S_1,S_2}(\mathbf{f}) \;\; = \;\; & K + p_{kk}f_k(1-f_k)N_k^2 \\ & + \textstyle\sum_{j\neq k} pf_jN_jN_k \\ & + p(N_+ - N_-)f_kN_k\end{aligned} \quad (6)$$

where we use N_+ and N_- to denote $\sum_{j\neq k}(1-f_j)N_j$, and $\sum_{j\neq k} f_jN_j$, respectively. Note that N_+ and N_- represent the size of the two partitions *excluding* object k.

If $(N_+ > N_-)$ then choosing $f_k = 0$ will give us a lower value for Cut_{S_1,S_2}, which also results in a partitioning vector whose $(\text{Size}_{S_2}(\mathbf{f}) - \text{Size}_{S_1}(\mathbf{f}))$ is larger than our starting vector. If $(N_+ < N_-)$ then choosing $f_k = 1$ will give us a lower value for Cut_{S_1,S_2}. In this case too the resulting partitioning vector would have $|\text{Size}_{S_2}(\mathbf{f}) - \text{Size}_{S_1}(\mathbf{f})|$ that is larger than for our starting vector. □

As a consequence of the above two theorems, we have

Corollary 2 *The possible candidates for the minimum cost partition are those vectors in S, Φ, and Ψ that have only one component that is different from all the others, i.e.* $\{\{\phi|\sum_i \phi_i = \frac{1}{N_i}\} \cup \{\psi|\sum_i \psi_i = 1 - \frac{1}{N_i}\} \cup \{\mathbf{s}|\sum_i s_i = 1\}\}$.

Theorem 7 *The recursive minimum cut strategy can be expected to result in correct partitioning if* $p < \{\frac{p_{11}}{N-N_1}, \cdots, \frac{p_{KK}}{N-N_k}\}$

Proof: Let $\phi_{\mathbf{i\neq 0}} \in \Phi$ be a vector whose all components are zero except for the i-th one. The expected cost of cut of such a partition would be $p_{ii}(N_i - 1) + p(N - N_i)$. Thus,

$$\text{Cut}_{S_1,S_2}(\phi_{i\neq 0}) = (p_{ii} - p)N_i + (pN - p_{ii}) \quad (7)$$

The minimum value of the above will be for the case when N_i is the minimum possible. Let $\mathbf{s_{i\neq 0}} \in S$ be vector whose all components are zero except for the i-th component. The expected cost of cut of such a partition would be

$$\text{Cut}_{S_1,S_2}(\mathbf{s_{i\neq 0}}) = pN_i(N - N_i) \quad (8)$$

The minimum value of the above will be for the case when N_i is the minimum possible.

Correct partitioning will result when a vector in S results in a lower cost partition than one in Φ or Ψ. Thus, the required condition is

$$\text{Cut}_{S_1,S_2}(\phi_{i\neq 0}) > \text{Cut}_{S_1,S_2}(\mathbf{s_{i\neq 0}}) \quad (9)$$

which can be transformed into

$$(p_{ii} - p(N - N_i))(N_i - 1) \quad > \quad 0 \tag{10}$$

The above will be always true if $p < \{\frac{p_{11}}{N-N_1}, \cdots, \frac{p_{KK}}{N-N_k}\}$. Since the vectors in Ψ represent the same partitions as the one in Φ, we do not need any other conditions. □

2.2. Average Cut

Theorem 8 *Minimum average cut cost is attained by partition vectors* $\mathbf{f} \in S$.

Proof: The matrix $\mathbf{P}$ can be expressed as sum of a diagonal and non-diagonal matrix: $\mathbf{P} = \mathbf{P_1} + \mathbf{P_2}$ where

$$\mathbf{P_1}(i,j) = \begin{cases} (p_{ii} - p)N_i^2 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \tag{11}$$

and

$$\mathbf{P_2}(i,j) = pN_iN_j \tag{12}$$

It is easy to see that $\mathbf{P_2} = p\mathbf{N}\mathbf{N}^T$. Thus, the expression for the expected average cut is given by

$$\begin{aligned} \text{AvgCut}_{S_1,S_2}(\mathbf{f}) &= N\frac{\mathbf{f}^T\mathbf{P}(\mathbf{1}-\mathbf{f})}{\mathbf{f}^T\mathbf{N}\mathbf{N}^T(\mathbf{1}-\mathbf{f})} \\ &= N\frac{\mathbf{f}^T\mathbf{P_1}(\mathbf{1}-\mathbf{f})}{\mathbf{f}^T\mathbf{N}\mathbf{N}^T(\mathbf{1}-\mathbf{f})} + pN \\ &= N\frac{\sum_i (p_{ii}-p)f_i(1-f_i)N_i^2}{\sum_i\sum_j f_i(1-f_j)N_iN_j} + pN \end{aligned} \tag{13}$$

Using Lemma 1 we can easily see that the first term on the right hand side is always positive and will attain a minimum value of 0 whenever all f_i's are all 0 or 1. The denominator of the first term is never 0 for valid partitions $\mathbf{f} \in S$. Thus, the minimum value of the average cut will, on an average, be attained by partition vectors in S.

2.3. Normalized Cut

We show the optimality of the Shi-Malik normalized cut measure by first showing that on an average it has a minimum at the correct partition for $K = 2$ objects and then we use proof by induction to extend it to the general K objects case. We also conjecture about the concave nature of the normalized cut measure, which we have not yet been able to prove or disprove.

Theorem 9 *For* $K = 2$ *objects, the normalized cut measure, on an average, attains its minimum value at* $f_1 = 0, f_2 = 1$ *or equivalently at* $f_1 = 1, f_2 = 0$.

Proof: The normalized cut cost for $K = 2$ can be expressed, using $x_1 = \frac{p_{11}N_1}{p_{12}N_2}$ and $x_2 = \frac{p_{22}N_2}{p_{12}N_1}$, in terms of $\text{NormCut}_{S_1,S_2}(0,1)$ as follows.

$$\begin{aligned} &\text{NormCut}_{S_1,S_2}(f_1,f_2) \\ &= \text{NormCut}_{S_1,S_2}(0,1) \\ &\times \frac{\begin{pmatrix} f_1(1-f_1)x_1 + f_2(1-f_2)x_2 \\ +(f_1(1-f_2) + f_2(1-f_1)) \end{pmatrix}}{\begin{pmatrix} f_1(1-f_1)\frac{x_1+1}{x_2+1} + f_2(1-f_2)\frac{x_2+1}{x_1+1} \\ +f_1(1-f_2) + f_2(1-f_1) \end{pmatrix}} \end{aligned} \tag{14}$$

Using Lemma 1, it is easy to show that (i) $x_1x_2 > 1$, (ii) $x_1 > \frac{x_1+1}{x_2+1}$, and (iii) $x_2 > \frac{x_2+1}{x_1+1}$. Using these inequalities it is easy to see that $f_1(1-f_1)x_1 + f_2(1-f_2)x_2 + (f_1(1-f_2) + f_2(1-f_1)) > f_1(1-f_1)\frac{x_1+1}{x_2+1} + f_2(1-f_2)\frac{x_2+1}{x_1+1} + f_1(1-f_2) + f_2(1-f_1)$. Hence, $\text{NormCut}_{S_1,S_2}(0,1)$ is always less than $\text{NormCut}_{S_1,S_2}(f_1 \neq 0, f_2 \neq 1)$. □

Theorem 10 *For* K *objects, the minimum value of the expected normalized cut cost is always attained by partition vectors* $\mathbf{f} \in S$.

Proof: We prove this by induction. As we saw in the previous theorem, this statement is true for $K = 2$. Let it be true for $K = M$. We show by contradiction that it has to be true for $K = M + 1$. Let us assume, that the minimum normalized cut cost partition for $M + 1$ objects is *not* one of the elements of the set S. An $(M + 1)$ objects case can be transformed into an M objects cases by conceptually merging any two objects, O_i and O_j into one object O_k, with the probability of feature association within the new object, p_{kk}, being presented by a value greater than p but smaller than either p_{ii} and p_{jj}. More precisely, with $p_{kk} = p + \frac{(p_{ii}-p)N_i^2 + (p_{jj}-p)N_j^2}{(N_i+N_j)^2}$. So, if $\mathbf{f_o}$ is an optimal solution for $M + 1$ objects case, it is also the optimal solution for one of the many possible transformed M objects cases. If $\mathbf{f_o} \notin S$, then it would suggest that we have an optimal solution for the M objects case that is not an element of S. This contradicts our assumption. □

Conjecture 1 *The expected value of the normalized cut measure is a concave function over the domain of all possible partitions.*

2.4. A Variation of the Normalized Cut

A variation of the Shi and Malik normalized partitioning measure is the cut weights normalized by the association within each partition, instead of the total valencies in each partition.

As we saw in Fig. 1, the shape of the expected value of this measure is not well formed. It is not a concave function, nor does it always have a minimum at $\mathbf{f} \in S$. Only for equal (nor nearly equal) sized objects the space over $\mathbf{f}$ is well behaved with the minimum at the desired partition. However, when one object is smaller than the other, $N_1 < N_2$, the surface is warped and the minimum is not at $f_1 = 0, f_2 = 1$.

Theorem 11 *The expected value of the variation of the normalized cut measure at* $\mathbf{f} = (0.5, \cdots, 0.5)^T$*, is 4 irrespective of the image statistics. This point is also an extremum.*

Proof: We can see this by taking the first derivative of the expected value of the measure with respect to $\mathbf{f}$ and computing the value at the extremum.

In the following theorem, we show that the variation of the normalized cut measure do not result in optimal partition for some image statistics for a $K = 2$ objects case, thus, diminishing its value as a good partitioning measure.

Theorem 12 *The point NormCutV2*$_{S_1,S_2}(0,1)$ *is* <u>*not*</u> *a global minimum for image statistics that satisfy*

$$\frac{p_{11}N_1}{pN_2} < 0.5 \ \text{ or } \ \frac{p_{22}N_2}{pN_1} < 0.5 \qquad (15)$$

Proof: In this proof, for notational simplicity, we will use x_1 to denote $\frac{p_{11}N_1}{pN_2}$ and x_2 to denote $\frac{p_{22}N_2}{pN_1}$. We find conditions under which NormCutV2$_{S_1,S_2}(0,1)$ is more than the value at NormCutV2$_{S_1,S_2}(0.5,0.5)$, which is always 4. It is easy to show by substitution in Eq. 5 that

$$\text{NormCutV2}_{S_1,S_2}(0,1) \;=\; \frac{2}{x_1}+\frac{2}{x_2} \qquad (16)$$

Using this and the fact that NormCutV2$_{S_1,S_2}(0.5,0.5) = 4$, we express our desired inequality as

$$\frac{1}{x_1}+\frac{1}{x_2} > 2 \qquad (17)$$

This would be true if $x_1 < 0.5$ or if $x_2 < 0.5$, which gives us the required conditions under which the variation of the normalized cut measure is not suitable. □

This condition can be true when the two objects have 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, p_{11}. For instance, if one object is 20% the size of the other and the association between feature within the smaller object is only twice as strong as that between the objects then the variation of the normalized cut measure cannot be expected to result in the correct partition.

3. Empirical Evaluation

The prediction from theory is that the behavior of the partition measures are different. The hypothesis of this empirical study is that the quality of the groups generated based on the three partition measures, i.e. min-cut, average cut, and normalized cut, are not sufficiently different to impact object recognition in a statistically significant manner. To this end, we consider the task of grouping constant curvature edge segments for object recognition by constrained search based techniques. We do not include the variation of the Shi-Malik normalized cut measure in this empirical study since, as we have seen, it is not well formed. We augment the min-cut strategy with additional logic to prevent splintering off small groups of features, which it is prone to do. We simply do not partition if size of one of the resulting partitions is smaller than an user specified minimum cluster size.

3.1. The Scene Structure Graph

For this study, we consider the grouping of extended image edge features, namely, constant curvature segments – straight line segments and circular arcs, which are represented as the nodes of the graph. The grouping framework is essentially the same as that described in [6].

The *Gestalt* inspired relationships of parallelism, perpendicularity (T and L-junctions), proximity, continuity, and common region form the basis for the formulation of the link weights between any two nodes representing the constant curvature segments. The links are quantified based on the attributes computed between any two edge segments such as min distance, max distance, slope difference, overlap and photometric attributes. Based on these photometric and geometric attributes, we classify and quantify the relationship between each pair of edge segment as being parallel, T–junction, L–junction, continuous, or "none–of–the–above". This is achieved using the maximum *aposteriori* probability (MAP) strategy based on the conditional probabilities of inferring the relationships based on the computed attributes. These priors and conditional probabilities are expressed in parametric forms whose parameters form the parameters of the grouping system. We also quantify the proximity and the common region factors between two edge segments in a MAP fashion. All the MAP reasoning is executed efficiently using the Bayesian network formalism. The sum of these three maximum *aposteriori* probabilities form the weights of the scene structure graph that is partitioned.

The average and the normalized cut measures based partitionings can be well approximated using graph spectral based partitioning strategies. We use the LEDA-4.0 implementation for the min-cut algorithm [8].

3.2. Performance Measure

We evaluate the quality of the final groups in terms of the expected performance of a constrained search based object recognition strategy starting from the groups. In [2], Borra and Sarkar proposed performance measures that do not require building a full vision system, but instead use manually specified ground truth. Let a grouping algorithm generate M groups $G_1,\ldots,G_M$. Let the number of objects in the scene be $O_1,\ldots,O_N$. For every pair of G_i and O_j that overlap, let (i) N_{G_i} denote the number of features in the detected group (G_i), (ii) N_{O_j} denote the number of model features (ground truth) in O_j, and (iii) $N_{G_i\cap O_j}$ denote the number of group features that lie on the model. Borra and Sarkar construct two performance measures that capture the expected time and quality of recognizing O_j based on group G_i.

$$P_{time}(G_i,O_j) = \frac{N_{G_i\cap O_j}}{N_{G_i}} \qquad (18)$$

$$P_{qual}(G_i,O_j) = \frac{N_{G_i\cap O_j}}{N_{O_j}} \qquad (19)$$

We then combine these two measures over all pairs of G_i and O_j that overlaps to get a composite performance given by,

$$\beta = \sqrt{\left(\frac{\sum_{ij} P_{time}(G_i, O_j)}{N_{overlaps}}\right)\left(\frac{\sum_{ij} P_{qual}(G_i, O_j)}{N_{overlaps}}\right)} \quad (20)$$

where, $N_{overlaps}$ denotes the total number of overlaps. This product form of combination of the two individual measures tends to assign equal importance to the time and quality of recognition. In addition, the used normalized summation form for each of the measures tends to (i) penalize a group that is spread across two objects more than a group that overlaps with one object and the background, (ii) prefer large groups over small groups, and (iii) penalizes groups of features that do not belong to any object.

Another possibility for performance measure are the false negative and false positive rates as measured by Amir and Lindenbaum [1], which are not directly tied to object recognition. We chose the option for task-based evaluation.

3.3. 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. Each of these images is associated with manual ground truth outlines of the objects of interest.

3.4. 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: 13 are used in the MAP based link quantification (Bayesian networks) that are used to quantify the graph links, 2 parameters are used by the graph spectral partitioning (minimum cluster size and maximum partition strength), 3 are edge detection parameters, 3 parameters are used by 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 [6]. With each parameter we associate one learning automaton that adaptively chooses from a set of possible actions (i.e the parameter values) on a random environment (i.e. the grouping algorithm and the images) so as to maximize the expected feedback (i.e. quality of the groups). Each automation used a probability vector, defined over the set of actions, to randomly choose actions (or parameter values). At start, every action is equally likely. At each iteration, based on the feedback from the environment and past action history, the action probability vector is updated so that the action with large feedback is more likely to be chosen than those that resulted in small feedback values. Over time, the action that maximizes the feedback will be chosen. It can be shown [4] that the team of learning automata is ϵ-optimal, i.e. the team will converge close to global optimal values. Out of 1500 to 5000 stochastic samplings, we select only those parameter sets with the 100 best performances on each of the 50 images for each of the 3 cut types.

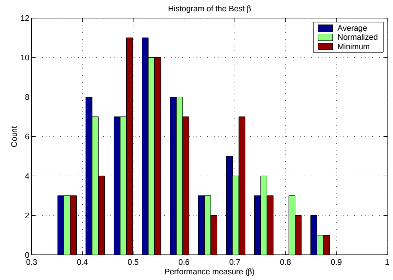


Figure 2. Histogram of the best βs (performance measure) for each of the cuts on all the 50 images.

3.5. Statistical Analysis

We present detailed evaluation of the grouping performances with the three graph partitioning measures. In the following subsections we first consider the following questions: Are the best performances for each cut type significantly different? Is the mean performance of the cut measures statistically different? Is the variation of performance for each cut type significantly different from each other? Is performance dependent on image class? We follow this by presentation of results for visual assessment and a discussion of time issues.

Best performances for each cut type: Do the performance for each partition cut measure significantly differ for the best parameter combination on each image? Fig 2 shows the histogram of the performances, β, for the best parameter combination on each of the images. We see that the histograms overlap and we cannot clearly separate out one from the other. This gross observation can be refined by analysis of variances (ANOVA) [5], which takes into account variations due to independent factors. The sources of performance variation are two: the image set and the cut type. Table 1 lists the results from analysis of variance, which shows that the best performances do not depend on the three cut measures used for grouping. In fact, we have found that the second best, third best, fourth best, and the fifth best performances also do not depend on the cut measure.

Source	F-value	P-value	Comments
Cut	2.08	0.1304	Not Significant
Image	35.81	0.0001	Significant

Table 1. ANOVA of the best performance on all images.

Mean performance of the cut measures: Here, we consider the mean performance of the 100 best performance, β_{mean}, on each image for each cut type. Are these mean performances statistically different? No. The results of the

ANOVA are shown in Table 2. We see, as before, that the cut type does not make a significant impact on the mean performance measure. We can also see from the F-value that the images are still a dominant source of variation.

Source	F-value	P-value	Comments
Cut	2.82	0.0644	Not Significant
Image	24.22	0.0001	Significant

Table 2. ANOVA of the $(\beta_1 + \beta_2 + .. + \beta_{100})/100$.

Variation of performance: Here we consider the variation in performance, which we quantify by the difference between the best performance and the 100^{th} best performance on each image, $\beta_{1-100} = \beta_1 - \beta_{100}$. Table 3 lists the ANOVA results. We can see that the range of performance β_{1-100} for the cut types are not significantly different.

Source	F-value	P-value	Comments
Cut	1.00	0.3732	Not Significant
Image	1.65	0.0180	Significant

Table 3. ANOVA of the $\beta_{1-100} = \beta_1 - \beta_{100}$.

Performance dependent on image class: As we have described above, we have 5 classes of images in the data set. Table 4 lists the mean values of the performances, for each image class and for each cut type. We observe clearly that for each of the image class there are not much variations in performance with respect to the cut measure, thus reinforcing our previous conclusions. We also observe that the performances on images of man made objects in indoor surroundings and aerial images are lower as compared to other classes.

To establish the statistical significances of our observation, we conducted ANOVA. The two independent factors, the cut measure and the image class, along with their interaction are the possible sources of variations. The results of ANOVA are shown in Table 5, where we see that, as before, the variation of performance due to the cut measure is not significant. In addition, the interaction between the image class and the cut type is also not significant. However, the variation due to the image classes is significant. Thus, the performance variations between images classes that we see in Table 4 are significant.

Visual assessment of results: Here, we present best grouping results on some images for visual assessment, which because of its subjective nature has to be done with caution. Visual assessment might not agree with the computed performance measure, β. In particular, we have to keep in mind that our performance measure (i) penalizes groups that straddles two objects more than groups that include part of an object and the background, and (ii) penalizes small groups.

Figs 3–5 shows the images on which performances of the cut measures differ the most. On Fig 3 (a) the average

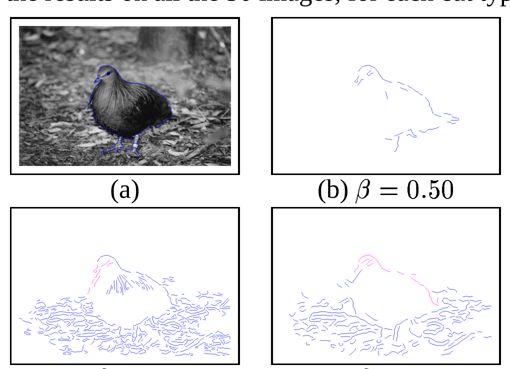
Image Class	Cut	Mean
1 (NO-I)	Average	0.60
	Normalized	0.61
	Minimum	0.61
2 (NO-O)	Average	0.61
	Normalized	0.61
	Minimum	0.62
3 (MO-I)	Average	0.49
	Normalized	0.50
	Minimum	0.53
4 (MO-O)	Average	0.60
	Normalized	0.65
	Minimum	0.64
5 (Aerial)	Average	0.50
	Normalized	0.48
	Minimum	0.47

Table 4. Mean values of the best performance index, β, for each class of images considered for each cut type. NO-I indicates natural objects indoors, NO-O indicates natural objects outdoors, MO-I indicates man–made objects indoors and MO-O indicates man–made objects outdoors.

Source	F-value	P-value	Comments
Cut	0.20	0.8184	Not Significant
Image Class	10.44	0.0001	Significant
Image Class*Cut	0.24	0.9820	Not significant

Table 5. ANOVA of the best performance on all images with interactions between the Image Class and the Cut type.

cut performs the best. Note that most of the background clutter is removed and the bird is grouped properly. The normalized cut is best for Fig 4 (a) and the min-cut is the best on Fig 5 (a). However, as we have seen earlier, this dependence of performance of each cut type on the image is not statistically significant. Variation due to image type swamps any variation due to the cut type. This reinforces the argument against the reliance on just visual assessment of results on a few images. On our website (`http://marathon.csee.usf.edu/PO.html`) we show the results on all the 50 images, for each cut type.



(a) (b) $\beta = 0.50$ (c) $\beta = 0.42$ (d) $\beta = 0.38$

Figure 3. Sample image on which average cut performance is most different from the others.(COLOR)

Time taken: We observed that the time taken to compute the normalized cut vary more than the other two cuts. This can be attributed to the normalized Laplacian matrix structure, all of whose diagonal entries are one and the rest of the entries are negative but with absolute values than one.

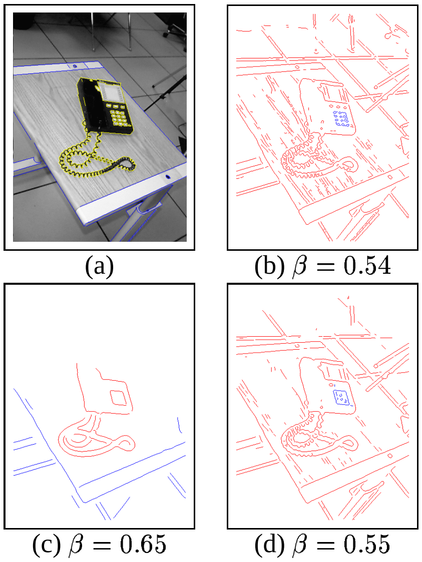


Figure 4. Sample image on which normalized cut performance is most different from the others.(COLOR)

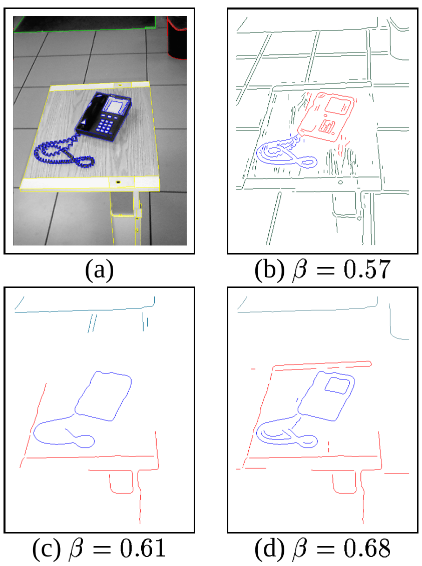


Figure 5. Sample image on which minimum cut performance is most different from the others.(COLOR)

Whereas for the average cut, the diagonal elements can have a value greater than 1. Hence the relative variation between the matrix element is less for the normalized cut case, which effects the rate of convergence of the eigenvector computations.

4. Conclusions and Implications

While the minimization of the average and the normalized cut measures, on an average, will result in correct partitioning, minimum cut and the variation of the normalized cut measures do not. The best, mean and the range of the 100 best performance, are not statistically significant with respect to the cut measure. The normalized cut based partition took significantly larger time to compute than the average and the mincut.

So, what are the implications of these conclusions that we have drawn from this study? First, the choice of the nature of the cut measure does not seem to be critically important for strategies that rely on grouping-by-graph-partitioning. The average cut measure seems to be somewhat a better choice from an theoretical optimality and time considerations. Second, importance of learning grouping parameters on a per image basis is obvious. Third, image type dominate as a source of variation in performance. We conjecture that the poor performance on certain image types such as aerial and man-made objects in indoor surroundings might be because the statistics of the objects of interest and the background are similar. Relations that are richer than just 2-ary ones might need to be considered when constructing the scene structure (hyper)-graph. Just 2-ary relations are probably inadequate for these domains.

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] D. W. Matula. Graph theoretic techniques for cluster analysis algorithms. In J. V. Ryzin, editor, *Classification and Clustering*, pages 95–129, 1977.

[4] K. S. Narendra and M. L. Thathachar. *Learning Automata: An Introduction*. Prentice Hall, 1989.

[5] T. Sanocki. *Student Friendly Statistics*. Prentice Hall, 2000.

[6] 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.

[7] 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.

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

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

[10] 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.

[11] 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.