A k-Partition, Graph Theoretic Approach to Perceptual Organization

Jeffrey Byrne, Avinash Gandhe, Ravi K. Prasanth, B. Ravichandan, Melvyn Huff and Raman K. Mehra,
Scientific Systems Company Inc., Woburn, MA, {jbyrne, avinash}@ssci.com
Sudeep Sarkar, *University of South Florida, Tampa, FL*
Sanjoy K. Mitter and Stefano Casadei, *Massachusetts Institute of Technology, Cambridge, MA*

Abstract—*This paper presents an k-partition, graph theoretic approach to perceptual organization. Principal results include a generalization of the bi-partition normalized cut to a k-partition measure, and a derivation of a sub-optimal, polynomial time solution to the NP-hard k-partition problem. The solution is obtained by first relaxing to an eigenvalue problem, followed by a heuristic procedure to enforce feasible solutions. This approach is a departure from the standard k-partitioning graph literature in that the partition measure used is non-quadratic, and is a departure from image segmentation literature in that k-partitioning is used in place of a recursive bi-partition. We apply this approach to image segmentation of infra-red (IR) images, and show representative segmentation results. Initial results show promise for further investigation.*

1 INTRODUCTION

Image segmentation is the process of extracting coherent regions or *object hypotheses* from images. This process is difficult since arbitrary lighting and backgrounds can clutter a scene, resulting in incorrect hypotheses. Segmentation systems often compensate for this variability by coupling bottom up processing, which creates salient regions and focuses computation, with top down knowledge based processing to overcome noise [1–3]. In this paper, we focus on a bottom up approach based on perceptual organization, which can be integrated into such a coupled system.

Perceptual organization refers to the ability of computer vision systems to organize features based on human vision, or Gestaltic, criteria. This sort of organization permits the formation of object hypotheses with minimal domain knowledge, and therefore minimal restrictions [4]. One efficient representation for perceptual organization is using graph theory [4, 5]. Images are abstracted to a graph representation, where graph nodes correspond to image locations, and weighted graph edges encode measures of Gestaltic similarity between nodes. This abstraction transforms the image segmentation problem into *graph partitioning*, which attempts to group nodes such that some partitioning measure or *cut* is maximized within each group and minimized between groups.

Approaches to graph partitioning can be classified according to partitioning measure and partitioning process. First, partitioning measures explored in image segmentation include the min cut, average cut, and normalized cut [6–8]. Analysis has shown that given 2-ary features, the average cut and normalized cut produce equivalent image segmentations which are superior to the min cut [9]. This analysis motivates further investigation into the normalized cut. Second, a partitioning process is generally categorized as a recursive bi-partition or k-partition [10]. A recursive bi-partition divides a graph into two groups, then recursively divides each group until k partitions are obtained [3, 5, 7]. A k-partition attempts to divide a graph into k groups simultaneously. Image segmentation has been historically dominated by the recursive bi-partition, however as noted in [11], recursive bi-partitioning suffers from a lack of global perspective which could lead to sub-optimal segmentation with respect to the k-partition.

In this paper, we present a k-partition graph theoretic approach to perceptual organization. We generalize the bi-partition normalized cut to a k-partition measure, and we derive a sub-optimal, polynomial time solution to the NP-hard k-partition problem. The solution is obtained by first relaxing to an eigenvalue problem, followed by a heuristic procedure to enforce feasible solutions. This approach is a departure from the standard k-partitioning graph literature in that the partition measure used is *non-quadratic* [12–15], and is a departure from image segmentation literature in that k-partitioning is used in place of a recursive bi-partition [3, 5, 7, 8]. We conclude with initial results from applying this approach to the segmentation of potential targets in IR images.

2 K-PARTITIONING A GRAPH

Image segmentation methods based on graph partitioning consist of formulating a weighted graph from the image and partitioning the graph so that some measure of similarity defined over partitions is minimized. Such minimization problems are in general NP-hard and, as a result, approximation

schemes are necessary.

Our approach has the following overall structure: *Formulate a convex lower bound problem and compute a suboptimal solution in polynomial-time using semidefinite programming (SDP) [12, 14, 16, 17].* We then define a lower bound using *eigenvalue-based bounds*, which are special cases of SDP that admit a non-iterative linear algebraic algorithm.

2.1 Notations

All vectors and matrices considered are real. $\mathbf{1}_k$ is a a column vector of length k whose elements are all equal to one. e_k is a column vector whose elements are all zeros except the kth entry which is 1. The transpose of A is denoted by A'. The element by element Hadamard product of two matrices A and B is denoted by $A \bullet B$. $A \geq B$ denotes element-wise inequality. $A \succeq B$ denotes the requirement that $A - B$ is positive semidefinite. Strict inequalities are denoted by $A > B$ and $A \succ B$. If $A \succeq 0$, then A has a positive square root which is denoted by $A^{1/2}$, i.e $A^{1/2} \succeq 0$ and $A = A^{1/2}A^{1/2}$. The element-wise positive square root of $A \geq 0$ is denoted by $\sqrt{A}$. $A^{\perp}$ an inner matrix whose columns span the null space of A'. $A \setminus B$ denotes the set difference such that $A \setminus B = \{x | x \in A, x \notin B\}$. The trace of a square matrix is denoted by $\mathrm{trace}\,(\cdot)$; while the maximum eigenvalue of a symmetric matrix is denoted by $\lambda_{\max}(\cdot)$. The $\mathrm{diag}\,(\cdot)$ operator acts on vectors to give diagonal matrices in the following way: (i,i)th element of $\mathrm{diag}\,(v)$ is the ith element of v.

2.2 Problem Overview

Let $G = (V, W)$ be a graph for which V is an ordered finite set of vertices or nodes with $|V| = N$ elements. Let $W = W'$ be the size $N \times N$ edge weighting matrix, such that $W \geq 0$, whose elements W_{ij} correspond to the similarity weight between the i^{th} and j^{th} nodes. The matrices

$$D = \mathrm{diag}\,(d) = \mathrm{diag}\,(W\mathbf{1}_N) \quad \text{and} \quad L = D - W$$

will play important roles in this paper. The diagonal elements of D are the sums of the rows of W, and we assume $D \succ 0$. The matrix L is known as the Laplacian, where $L\mathbf{1}_N = 0$, $\det(L) = 0$, and $L \succeq 0$.

A k-partition of G is an $N \times K$ membership matrix X, which assigns each $v \in V$ to one partition V_i such that for all $1 \leq i \leq k$, $V_i \neq \emptyset$, $V_1 \cup V_2 \cup \cdots \cup V_k = V$, and $V_i \cap V_j = \emptyset$ for $i \neq j$. An optimal k-partition X_{opt} minimizes a similarity measure $\mathcal{N}(X)$ over all X satistfying certain conditions. The goal of this analysis is to define an optimal k-partition X_{opt} of G.

2.3 Similarity measures

We are interested in separating the vertices in V into k disjoint classes according to some measure of similarity. Define:

$$\begin{aligned} \mathcal{V}_p = \quad & \{(A_i)_{i=1}^k : A_i \neq \emptyset,\ A_i \cap A_j = \emptyset \text{ for all } i \neq j, \\ & \text{and } A_1 \cup A_2 \cup \cdots \cup A_k = V\} \end{aligned} \tag{2}$$

to be the universe of partitions over which similarity measures are defined. In order to define similarity measures and perform computations, let us first identify $\mathcal{V}_p$ with the following subset of a vector space:

$$\begin{aligned} \mathcal{X}_p = \quad & \{X \in \mathbb{R}^{N \times k} : X \bullet X = X,\ X\mathbf{1}_k = \mathbf{1}_N \text{ and} \\ & X'\mathbf{1}_N \geq 1\} \end{aligned} \tag{3}$$

This is a finite set because the constraint $X \bullet X = X$ implies that every element of X must be either 0 or 1. In addition, the constraints $X'\mathbf{1}_N \geq 1$ (non-empty sets) and $X\mathbf{1}_k = \mathbf{1}_N$ (disjoint sets) guarantee that the columns are orthogonal and that one and only one entry of each row is 1. It is easy to verify that $\mathcal{V}_p$ and $\mathcal{X}_p$ are in one-to-one onto correspondence.

Normalized cut— Define the functions $f_p, g_p : \mathcal{X}_p \to \mathbb{R}^{k \times k}$ as follows:

$$f_p(X) = X'DX \ \text{ and } \ g_p(X) = X'LX \tag{4}$$

where D is the diagonal matrix of row sums of W and L is the Laplacian. Some properties of these functions are given below:

1. f_p is diagonal because D is diagonal, the columns of X are orthogonal and every element of X is either 0 or 1. Moreover, $f_p(X) \succ 0$ for any X, a consequence of $D \succ 0$ and the fact that X has full rank.

2. $g_p(X) \succeq 0$, $g_p(X)\mathbf{1}_k = 0$ and $\det(g_p(X)) = 0$ for any X. These follow from the properties of L and X.

Before defining the measures, let us interpret these functions in terms of the set of partitions $\mathcal{V}_p$. Fix $X \in \mathcal{X}_p$. Let $(A_i)_{i=1}^k$ be the corresponding partition of V. The (i,i)th element of $f_p(X)$ is the total connection weight from A_i to V. Notice that

$$g_p(X) = X'LX = X'(D - W)X = f_p(X) - X'WX$$

So, the (i,i)th element of $g_p(X)$ is the total connection weight from A_i to $V \setminus A_i$. Therefore, the (i,i)th element of

$$f_p(X)^{-1/2} g_p(X) f_p(X)^{-1/2}$$

is the total connection weight from A_i to $V \setminus A_i$ normalized by the total connection weight from A_i to V. With this in mind, we define the *normalized cut* $\mathcal{N}_0 : \mathcal{X}_p \to \mathbb{R}$ as follows:

$$\mathcal{N}_0(X) = \mathrm{trace}\left(f_p(X)^{-1/2} g_p(X) f_p(X)^{-1/2}\right) \tag{5}$$

This is clearly a generalization of the normalized cut of Shi and Malik [7].

Roughly speaking, $\mathcal{N}_0$ sees each class A_i in relation to its complement in V and does not directly use *between-class* information as the definition involves only diagonal elements. The off-diagonal elements of $-g_p(X)$ are the total weights between classes and could be useful in separating the classes. One way to bring in this *between-class* information is to define the *maximum eigenvalue measure* $\mathcal{N}_\lambda : \mathcal{X}_p \to \mathbb{R}$:

$$\mathcal{N}_\lambda(X) = \lambda_{\max}\left(f_p(X)^{-1/2} g_p(X) f_p(X)^{-1/2}\right) \quad (6)$$

Interestingly, $\mathcal{N}_0 = \mathcal{N}_\lambda$ in the 2-cut case ($k = 2$) which is to be expected from our interpretations. Mathematically, this is because one of the eigenvalues of $f_p^{-1/2} g_p f_p^{-1/2}$ is always zero.

2.4 Graph Partitioning Problem

We would like to minimize the measure $\mathcal{N}_\lambda$ defined in the previous section possibly over $\mathcal{X}_p$. However, $\mathcal{X}_p$ contains all k-cuts including some that partition the image into several small subsets and a large one. The normalized cut measure was originally defined by Shi and Malik [7] to guard against such partitions from becoming optimal. *A more direct way is to insist that every subset in a partition be bigger than a specified size thereby eliminating unrealistic partitions*. To this end, let $s_0 \in \mathbb{R}^k$ be a vector whose entries are all strictly positive integers that specify the minimum size of each partition. Assume that the sum of elements of s_0 is no larger than N. Define:

$$\begin{aligned}\mathcal{S}_0 = \quad & \{s \geq s_0 : \text{entries of } s \text{ are integers and} \\ & \mathbf{1}_k{}' s = N\}\end{aligned} \quad (7)$$

and, for each $s \in \mathcal{S}_0$, define the following subset of $\mathcal{X}_p$:

$$\begin{aligned}\mathcal{X}_p^s = \quad & \{X \in \mathbb{R}^{N\times k} : X \bullet X = X,\ X\mathbf{1}_k = \mathbf{1}_N \\ & \text{and } X'\mathbf{1}_N = s\}\end{aligned} \quad (8)$$

which is the set of all partitions of X whose components are of size specified by the components of s. The union of $\mathcal{X}_p^s$ over $\mathcal{S}_0$ is a more reasonable representative of image segmentation objectives than $\mathcal{X}_p$. We shall therefore formulate graph partitioning problems on this union.

The goal of the graph partitioning problem is to find a sub-optimal solution ($s \in \mathcal{S}_0, X \in \mathcal{X}_p^s$) such that $\mathcal{N}_\lambda(X) < \gamma_\lambda + \epsilon$ for any $\epsilon > 0$, where

$$\gamma_\lambda = \min_{s\in\mathcal{S}_0} \min_{X\in\mathcal{X}_p^s} \mathcal{N}_\lambda(X) \quad (9)$$

It is known that the normalized 2-cut is NP-complete. So, polynomial-time algorithms for computing γ_λ cannot be found (unless P=NP). We are therefore interested in computing sub-optimal solutions preferably with known quality.

2.5 Eigenvalue-based bounds

The basic idea in semidefinite programming (SDP) relaxation and its special case considered in this section is to replace the constraint set $\mathcal{X}_p^s$ with a larger set containing $\mathcal{X}_p^s$. It can be shown [18] that $\mathcal{X}_p^s$ in (8) is equivalent to:

$$\begin{aligned}\mathcal{X}_p^s = \quad & \{X \in \mathbb{R}^{N\times k} : X'X = \mathrm{diag}\,(s),\ X\mathbf{1}_k = \mathbf{1}_N, \\ & X'\mathbf{1}_N = s \text{ and } X \geq 0\}\end{aligned} \quad (10)$$

from which a relaxation is obtained by ignoring the constraint $X \geq 0$.

Define:

$$\begin{aligned}\mathcal{X}_{\text{relax}}^s = \quad & \{X \in \mathbb{R}^{N\times k} : X'X = \mathrm{diag}\,(s), \\ & X\mathbf{1}_k = \mathbf{1}_N \text{ and } X'\mathbf{1}_N = s\}\end{aligned} \quad (11)$$

which, unlike $\mathcal{X}_p^s$, is an infinite set, due to the relaxation.

It can be shown [18] that $X \in \mathcal{X}_{\text{relax}}^s$ if and only if

$$X(s) = \begin{bmatrix}\frac{\mathbf{1}_N}{\sqrt{N}} & \mathbf{1}_N{}^\perp\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & Z\end{bmatrix}\begin{bmatrix}\frac{\sqrt{s}'}{\sqrt{N}}\\ \sqrt{s}^{\perp\prime}\end{bmatrix}\mathrm{diag}\,(s)^{1/2} \quad (12)$$

or correspondingly, $X(s) = U_1 U_z U_s' \mathrm{diag}\,(s)^{1/2}$ for some inner matrix Z.

Using this definition for $\mathcal{X}_{\text{relax}}^s$, a lower bound to the minimization in (9) for a fixed $s \in \mathcal{S}_0$ is:

$$\begin{aligned}\inf_{\gamma,Z}\ \gamma \quad & \text{subject to } X'(\gamma D - L)X \succ 0, \\ & X = U_1 U_z U_s' \mathrm{diag}\,(s)^{1/2},\ Z'Z = I\end{aligned} \quad (13)$$

which is obtained by introducing a new variable γ and using properties of $\lambda_{\max}(\cdot)$. Removing X and the invertible outer factors using congruence gives:

$$\begin{aligned}\gamma_{\text{lbd}}^\lambda = \quad & \inf_{\gamma,Z}\ \gamma \ \text{ subject to } U_z' U_1' (\gamma D - L) U_1 U_z \succ 0 \\ & \text{and } Z'Z = I\end{aligned} \quad (14)$$

Thus, the inner minimization of the lower bound problem in (9) is equivalent to the above infimization.

Using a Schur complement argument, the positive definite constraint $U_z' U_1' (\gamma D - L) U_1 U_z \succ 0$ in (14) is equivalent to the constraint:

$$Z'\left[(\gamma - 1) I + \sqrt{d}^{\perp\prime} D^{-1/2} W D^{-1/2} \sqrt{d}^{\perp}\right] Z \succ 0 \quad (15)$$

Define:

$$A = (\gamma - 1) I + \sqrt{d}^{\perp\prime} D^{-1/2} W D^{-1/2} \sqrt{d}^{\perp}$$

$$B = D^{-1/2} W D^{-1/2}$$

The positive definite constraint in (15) holds if Z is any matrix whose range is contained in the eigen subspace associated with the strictly positive eigenvalues of A. It follows from a similarity transform argument that if the eigenvalues of B are $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, then the eigenvalues of A are $(\gamma - 1) + \lambda_2 \geq (\gamma - 1) + \lambda_3 \geq \cdots \geq (\gamma - 1) + \lambda_N$. For A to have $k - 1$ strictly positive eigenvalues, the requirement $\gamma > (1 - \lambda_k)$ must hold. Therefore, the choice $\gamma = (1 - \lambda_k) + \epsilon$ for some $\epsilon > 0$, and Z an inner matrix whose columns are the orthonormal eigenvectors corresponding to the k-1 strictly positive eigenvalues of A, results in the pair (γ, Z) being a feasible, sub-optimal solution to (13) and (14). For a given $s \in \mathcal{S}_0$, this feasible solution defines an X as in (12), for which $\mathcal{N}_\lambda(X) < (1 - \lambda_k) + \epsilon$. For a detailed derivation, please see [18].

2.6 Computational Solution

Following the results in the previous section, a sub-optimal solution $(s \in \mathcal{S}_0, X \in \mathcal{X}^s_{\text{relax}})$ such that $\mathcal{N}_\lambda(X) < \gamma^\lambda_{\text{lbd}} + \epsilon$ given $\epsilon > 0$, can be calculated as follows.

Algorithm 2.1 (Solution of eigenvalue-based lower bound) Perform the following computations:

1. Compute largest k eigenvalues of $D^{-1/2} W D^{-1/2}$
$$1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$$
and a corresponding orthonormal set of eigenvectors $v_1, v_2, \cdots, v_k$.

2. Define
$$q_i = D^{-1/2} v_i \quad \text{for } i = 1, 2, \cdots, k$$
and apply Gram-Schmidt orthonormalization procedure starting with q_1.

3. Choose any $s \in \mathcal{S}_0$ and set
$$X = M \begin{bmatrix} \sqrt{s}' / \sqrt{N} \\ \sqrt{s}^{\perp'} \end{bmatrix} \text{diag}\,(s)^{1/2}$$
where $M \in \mathbb{R}^{N \times k}$ is the inner matrix obtained in the Gram-Schmidt orthogonalization of Step 2.

Then, $\gamma^\lambda_{\text{lbd}} = 1 - \lambda_k$ and, for any $\epsilon > 0$, the matrix X given in Step 3 satisfies $\mathcal{N}_\lambda(X) < 1 - \lambda_k + \epsilon$.

Note that Step 1 and Step 2 of the algorithm do not involve s. Hence, $\gamma^\lambda_{\text{lbd}}$ is independent of s and every X given by Step 3 has the same quality. This is because the relaxed constraint set, without the cone constraint $X \geq 0$, admits all directions. So, in Step 3, we ought to choose $s \in \mathcal{S}_0$ with the objective of enforcing $X \geq 0$. We do so by solving the following minimization problem for $\widehat{s}$:

$$J = \min_{\substack{\widehat{s} \geq \widehat{\mathcal{S}}_0 \\ K = -K^T}} ||X(\widehat{s}) \bullet X(\widehat{s}) - X(\widehat{s})||^2 + ||X(\widehat{s})^-||$$

where
$$X(\widehat{s}) = M \begin{bmatrix} \sqrt{\widehat{s}'} / \sqrt{N} \\ e^K (\sqrt{\widehat{s}}^{\perp'}) \end{bmatrix} \text{diag}\,(\widehat{s})^{1/2}$$

$X(\widehat{s})^-$ is the matrix $X(\widehat{s})$ with all positive entries set to 0, $||\bullet||$ denotes the Frobenius matrix norm, and:

$$\widehat{\mathcal{S}}_0 = \{\widehat{s} : \widehat{s} \geq s_0 \text{ and } \mathbf{1}_k{}' \widehat{s} = N\}$$

which is a relaxation of $\mathcal{S}_0$ defined in (7).

The final result X is formed as in Step 3 of Algorithm 2.1 using the solution $s = \widehat{s}$ to the minimization of J. The matrix X defines a k-partition, where the rows of the solution can be interpreted as the confidence that a node belongs the group represented by the corresponding column of X. Due to the use of relaxed constraints, X is real valued, so the group membership of the i^{th} node in X is chosen to be j such that $X_{ij} \geq X_{il}$ for all $1 \leq l \leq k$.

3 IMPLEMENTATION AND RESULTS

Section 2 described a computational engine to perform a simultaneous k-way partitioning of a weighted graph. In this section, we relate the theory of Section 2 to a specific application, namely, segmenting objects of interest from background in IR imagery.

Image segmentation can be considered to be a preliminary step towards Automatic Target Detection (ATD), the autonomous location of targets in sensor data. When using infrared and optical sensors, where the sensor data is represented by an image of the scene under surveillance, the ATD problem becomes one of forming target-class hypothesis from pixels in the image. Image segmentation is an effective pre-processing tool to form object hypotheses before a decision level algorithm is used to determine if each of these objects is a target of interest. The use of perceptual organization ensures that the object hypotheses are formed based on robust, human vision based criteria.

3.1 Features for grouping

Perceptual organization theory provides a set of heuristics believed to be used by the human visual system to separate figure from ground in an image, with little or no prior information. These heuristics grew out of the Gestaltic movement to explain how humans are able to reliably and repeatedly segment arbitrary images into figure and background. Such heuristics include: proximity, similarity, closure, symmetry and continuity [4].

In the context of graph partitioning, perceptual organization can be used to define features which characterize the relationship between image regions. The relationships are encoded in the edge weights of the weighted graph.

The weight between two nodes in the graph, or in our partic-

ular case, two points in the image, is calculated as a function of image properties, or *features* at the two points. Specifically, with M features under consideration, the weight W_{ij} between two pixels, i and j, can be calculated as

$$W_{ij} = W_{ij}^{feature_1} \times W_{ij}^{feature_2} \times ... \times W_{ij}^{feature_M} \quad (16)$$

The weight based on a certain feature is a function of some measure of similarity, d_{ij}, such that,

$$W_{ij}^{feature_m} = exp\left[-\frac{(d_{ij}^{feature_m})^2}{\sigma_{feature_m}^2}\right] \quad (17)$$

where, $\sigma_{feature_m}^2$, determines the relative strength of the feature in the perceptual grouping process.

Features explored in this implementation are as follows:

- *Intensity:* $d_{ij}^{intensity}$ is the difference in the intensity of the image at pixels i and j.
- *Proximity:* $d_{ij}^{proximity}$ is the spatial distance between pixels i and j.
- *Texture:* $d_{ij}^{texture}$ is a measure of the difference in distributions of *textons*, or primitive texture elements within a region. As in [19], we calculate the similarity between two pixels to be the χ^2 distance between two texton histograms computed from local regions centered on each pixel.
- *Contour:* $d_{ij}^{contour}$ is calculated using the *intervening contour* framework [19], where the similarity between the two pixels is inversely proportional to the strongest contour that separates them.

3.2 Computational Considerations

The computational requirements of the k-way graph partitioning algorithm increase polynomially with image size [20]. Any effort to decrease the computational requirements by downsampling the image or weight matrix results in small targets being overlooked, which implies that it is important to use every pixel on target, and by extension, every pixel in the image, for grouping.

A possible solution to the computational problems encountered when dealing with every pixel in a large image is to develop a *grouping hierarchy*. We have implemented such a hierarchy using a *bottom up* approach where small areas of the image are partitioned at the lowest level of the hierarchy using the approach described in Section 2 and these partitions are then recursively merged at the higher levels. The recursive merge also uses the theory in Section 2 by treating the partitions as nodes in a graph and using a measure of similarity between partitions as edge weights. We illustrate our approach using the example shown in Figure 1. For details, one may refer to [20]. By using a quadrant-based recursive merging process, this approach is approximately 16^{L-1} times faster than a single-step partitioning of the entire image at once, where L is the number of levels in the hierarchy.

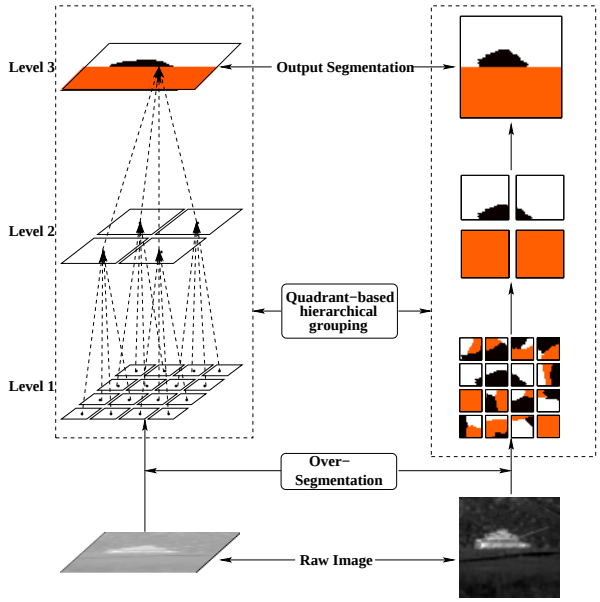
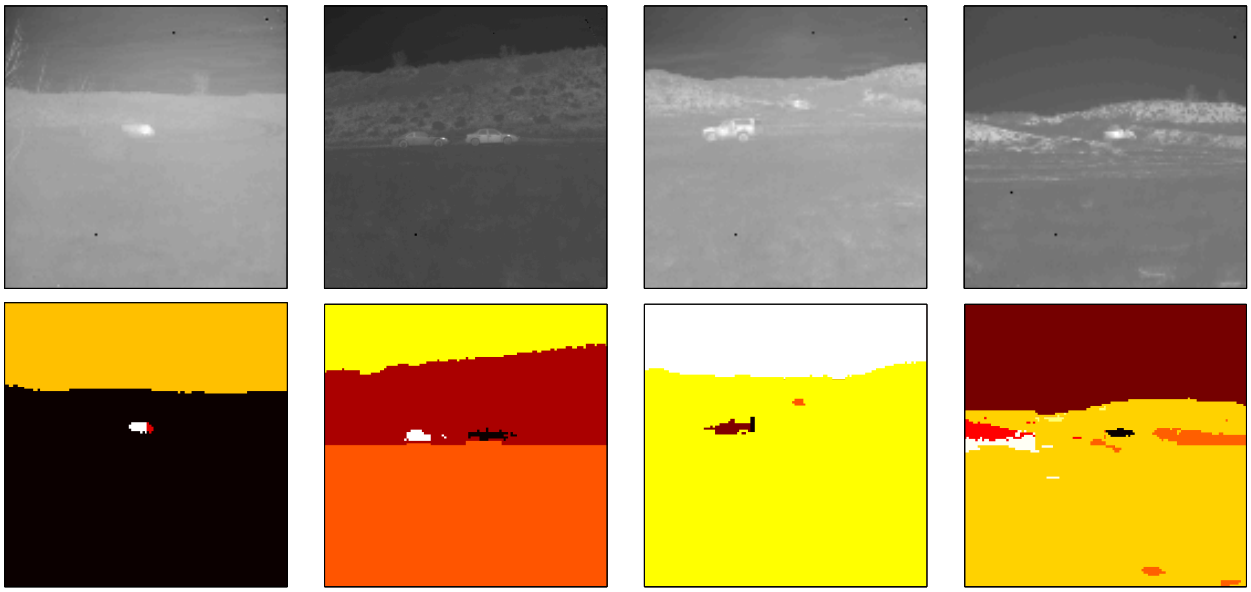


Figure 1. The hierarchical approach for a 3-level hierarchy

3.3 Results

This section shows the results of the following approach: creating the image graph as defined in Section 3.1, partitioning the graph using the theory described in Section 2 and then repeating these steps using the recursive hierarchy described in Section 3.2. The images shown here are from the Fort Carson database [21], downsampled from the original size of 256×256 to 128×128. The results are shown with different colors, each corresponding to a perceptually salient image region. It should be noted that we currently fix the number of partitions, k, of a sub-image, at each level of the hierarchy.

Figure 2 shows examples of our approach, with the IR images under test along the top row and the corresponding segmentation results along the bottom row. Figure 2 (a) shows a simple example, where the target of interest is an M113 TOW vehicle, against the backdrop of a grassy region. The result shows that the three perceptually salient regions of grass, sky and target are correctly segmented. Figure 2 (b) shows an image with two targets of interest. The car on the right is clearly visible while the car on the left bleeds slightly into the background. Both targets are successfully segmented along with the correct background partitions between the ground, a hilly area and the sky. Figure 2 (c) shows an IR image with two targets of interest, an SUV that is clearly visible and distant M113 TOW vehicle, along the slope of a hill, that is very hard to see. The algorithm detects the SUV and partially detects the M113, along with correctly identifying the horizon. Figure 2 (d) shows an M113 TOW vehicle against a cluttered background in the IR image. The segmentation algorithm correctly segments the target, indicated by the black region in the result, from background, along with additional clutter.



(a) (b) (c) (d)

Figure 2. (top) A series of images from the Fort Carson IR imagery database (bottom) Segmentation results

4 CONCLUSIONS

This paper has presented the theory and initial results for a new k-way graph partitioning approach to image segmentation based on perceptual organization. The results in Figure 2 show that the approach provides good target-background separation in scenarios with multiple targets, targets of different sizes and cluttered background. In all these cases, the background is also segmented appropriately.

The results shown in Figure 2 also point out areas that need further investigation. For instance, in Figures 2(a) and 2(c), the targets are oversegmented. Such results show the importance of automatically selecting the appropriate k for partitioning the image. Also, one can see from Figure 2(b) that the car on the left is only detected partially. Given the features we are using, this segmentation is acceptable, since the IR image shows that the bottom of the car does blend into the background. However, a human is able to segment the whole car by making use of human vision cues such as continuity and familiarity, which implies that target detection may be improved by introducing additional Gestaltic features. Finally, Figure 2(d) shows detection of clutter along with the target. This highlights the need for a coupled top down processing to be integrated with bottom up processing to reduce noise.

In summary, we believe the results shown are promising in the context of target detection, and conclude from these observations and results that image segmentation based on perceptual organization merits further investigation. Future work includes a quantitative analysis of this approach vs. current segmentation algorithms and other graph partitioning approaches to show comparative results, further tests to show robustness given scene variability, and investigation into methods to automatically determine an appropriate k.

ACKNOWLEDGMENTS

This work was performed with support under Army Research Office (ARO) contracts DAAD19-00-C-0060 and DAAD19-01-C-0057. We would like to thank Dr. Linda Bushnell, Dr. Hua Wang, and Dr. William Sander for their support during the course of this work.

REFERENCES

[1] P. Halvadar, G. Medioni, and F. Stein, "Perceptual grouping for generic recognition," *International Journal of Computer Vision*, **20**, *59-80*, 1996.

[2] R. C. Nelson and A. Selinger, "A perceptual grouping hierarchy for 3D object recognition and representation," *DARPA Image Understanding Workshop*, *157-163*, 1998.

[3] S. Yu, R. Gross and J.Shi, "Concurrent object recognition and segmentation by graph partitioning," *Neural Information Processing Systems*, Dec 9-14, 2002.

[4] S. Sarkar and K. L. Boyer, "Perceptual organization in computer vision: A review and proposal for a classificatory structure," *IEEE trans. on Systems, Man, and Cybernetics*, **23(2)**, *382-399*, 1993.

[5] Y. Weiss, "Segmentation using eigenvectors: A unified view," *Proceedings of the International Conference on Computer Vision*, **2**, *975-982*, 1999.

[6] S. Sarkar and P. Soundararajan, "Supervised learning of large perceptual organization: Graph spectral partition-

ing and learning automata," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22(5)**, *504-525*, May 2000.

[7] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22(8)**, *888–905*, August 2000.

[8] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and application to image segmentation," *IEEE trans. on Pattern Analysis and Machine Intelligence*, **15**, *1101-1113*, November, 1993.

[9] P. Soundararajan and S. Sarkar, "Investigations of measures for grouping by graph partitioning," *International Conference on Computer Vision and Pattern Recognition*, December 2001.

[10] U. Elsner, "Graph partitioning: A survey," *Technical Report SFB393/97-27*, TU Chemnitz, 1997.

[11] H.D. Simon and S-H. Teng, "How good is recursive bisection?," *SIAM J. Scientific Computing*, **18(5)**, *1436-1445*, 1997.

[12] F. Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," *SIAM J. of optimization*, **5**, *13-51*, 1995.

[13] J. Falkner, F. Rendl and H. Wolkowicz, "A computational study of graph partitioning," *Mathematical Programming*, **66**, *211-240*, 1994.

[14] L. Lovász and A. Schrijver, "Cones of matrices and setfunctions, and 0-1 optimization," *SIAM J. of optimization*, **1**, *166-190*, 1991.

[15] H. Wolkowicz and Q. Zhao, "Semidefinite programming relaxations for the graph partitioning problem", 1999.

[16] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial methods in convex programming*, SIAM, Philadelphia, 1994.

[17] L. Vandenberghe and S. Boyd, "Semidefinite programming,"*SIAM Review*, **38**, *49-95*, 1996.

[18] R. K. Prasanth, A. Gandhe, B. Ravichandran, S. Casadei, S. Mitter and S. Sarkar, "ATR using perceptual organization" *Tech. Report, SSCI-1297-interim*, Scientific Systems Company Inc., Woburn, MA, April 2002.

[19] J. Malik, S. Belongie, T. K Leung and J. Shi, "Contour and texture analysis for image segmentation," *International Journal of Computer Vision*, **43(1)**, *7–27*, 2001.

[20] A. Gandhe, J. Byrne, M. Huff and B. Ravichandran, "ATR using perceptual organization, " *Tech. Report SSCI-1297-8* Scientific Systems Company Inc., Woburn, MA, May 2003.

[21] J. Beveridge Ross, D. Panda, and T. Yachik, "Fort Carson RSTA data collection," *Final Report CS-94-118*, Colorado State University, November 1993.