Normalized Cuts Revisited:
A Reformulation for Segmentation with Linear Grouping Constraints

Anders P. Eriksson
anderspe@maths.lth.se

Carl Olsson
calle@maths.lth.se

Fredrik Kahl
fredrik@maths.lth.se

Centre for Mathematical Sciences
Lund University, Sweden

Abstract

Indisputably Normalized Cuts is one of the most popular segmentation algorithms in computer vision. It has been applied to a wide range of segmentation tasks with great success. A number of extensions to this approach have also been proposed, ones that can deal with multiple classes or that can incorporate a priori information in the form of grouping constraints. However, what is common for all these suggested methods is that they are noticeably limited and can only address segmentation problems on a very specific form. In this paper, we present a reformulation of Normalized Cut segmentation that in a unified way can handle all types of linear equality constraints for an arbitrary number of classes. This is done by restating the problem and showing how linear constraints can be enforced exactly through duality. This allows us to add group priors, for example, that certain pixels should belong to a given class. In addition, it provides a principled way to perform multi-class segmentation for tasks like interactive segmentation. The method has been tested on real data with convincing results.

1. Image Segmentation

Image segmentation can be defined as the task of partitioning an image into disjoint sets. This visual grouping process is typically based on low-level cues such as intensity, homogeneity or image contours. Existing approaches include thresholding techniques, edge based methods and region-based methods. Extensions to this process includes the incorporation of grouping constraints into the segmentation process. For instance the class labels for certain pixels might be supplied beforehand, through user interaction or some completely automated process, [8, 2].

Currently the most successful and popular approaches for segmenting images are based on graph cuts. Here the images are converted into undirected graphs with edge weights between the pixels corresponding to some measure of similarity. The ambition is that partitioning such a graph will preserve some of the spatial structure of the image itself. These graph methods based were made popular first through the Normalized Cut formulation of [9] and more recently by the energy minimization method of [3]. This algorithm for optimizing objective functions that are submodular has the property of solving many discrete problems exactly. However, not all segmentation problems can be formulated with submodular objective functions, nor is it possible to incorporate all types of linear constraints.

The work described here concerns the former approach, Normalized Cuts, the relevance of linear grouping constraints and how they can be included in this framework. It is not the aim of this paper to argue the merits of one method, or cut metric, over another, nor do we here concern ourselves with how the actual grouping constraints are obtained. Instead we will show how through Lagrangian relaxation one in a unified can handle such linear constrains and also in what way they influence the resulting segmentation.

1.1. Problem Formulation

Consider an undirected graph $\mathbf{G}$, with nodes $\mathbf{V}$ and edges $\mathbf{E}$ and where the non-negative weights of each such edge is represented by an affinity matrix W, with only non-negative entries and of full rank. A min-cut is the non-trivial subset A of V such that the sum of edges between nodes in A and its complement is minimized, that is the minimizer of

$$cut(A, V) = \sum_{\substack{i \in A \\ j \in V \setminus A}} w_{ij} \tag{1}$$

This is perhaps the most commonly used method for splitting graphs and is a well known problem for which very efficient solvers exist. It has however been observed that this criterion has a tendency to produced unbalanced cuts, smaller partitions are preferred to larger ones.

In an attempt to remedy this shortcoming, Normalized Cuts was introduced by [9]. It is basically an altered criterion for partitioning graphs, applied to the problem of perceptual grouping in computer vision. By introducing a normalizing term into the cut metric the bias towards undersized cuts is avoided. The Normalized Cut of a graph is defined as:

$$N_{cut} = \frac{cut(A,V)}{assoc(A,V)} + \frac{cut(B,V)}{assoc(B,V)} \tag{2}$$

where $A \cup B = V$, $A \cap B = \emptyset$ and the normalizing term defined as $assoc(A,V) = \sum_{i \in A, j \in V} w_{ij}$. It is then shown in [9] that by relaxing (2) a continuous underestimator of the Normalized Cut can be efficiently computed. These techniques are then extended in [11] beyond graph bipartitioning to include multiple segments, and even further in [12] to handle certain types of linear equality constraints.

One can argue that the drawbacks of this, the classical formulation, for solving the Normalized Cut are that firstly obtaining a discrete solution from the relaxed one can be problematic. Especially in multiclass segmentation where the relaxed solution is not unique but consists of an entire subspace. Furthermore, the set of grouping constraints is also very limited, only homogeneous linear equality constraints can be included in the existing theory. We will show that this excludes many visually relevant constraints. In [4] an attempt is made at solving a similar problem with general linear constraints. This approach does however effectively involve dropping any discrete constraint all together, leaving one to question the quality of the obtained solution.

2. Normalized Cuts with Grouping Constraints

In this section we propose a reformulation of the relaxation of Normalized Cuts that in a unified way can handle all types of linear equality constraints for any number of partitions. First we show how we through duality theory reach the suggested relaxation. The following two sections then show why this formulation is well suited for dealing with general linear constraints and how this proposed approach can be applied to multiclass segmentation.

Starting off with (2), the definition of Normalized Cuts, the cost of partitioning an image with affinity matrix W into two disjoint sets, A and B, can be written as

$$N_{cut} = \frac{\sum_{\substack{i \in A \\ j \in B}} w_{ij}}{\sum_{\substack{i \in A \\ j \in V}} w_{ij}} + \frac{\sum_{\substack{i \in B \\ j \in A}} w_{ij}}{\sum_{\substack{i \in B \\ j \in V}} w_{ij}}. \tag{3}$$

Let $z \in \{-1,1\}^n$ be the class label vector, W the $n \times n$-matrix with entries w_{ij}, d the $n \times 1$-vector containing the row sums of W, and D the diagonal $n \times n$-matrix with d on the diagonal. A 1 is used to denote vectors of all ones. We can write (3) as

$$\begin{aligned} N_{cut} &= \tfrac{\sum_{i,j} w_{ij}(z_i - z_j)^2}{2\sum_i (z_i+1)d_i} + \tfrac{\sum_{i,j} w_{ij}(z_i - z_j)^2}{2\sum_i (z_i-1)d_i} = \\ &= \tfrac{z^T(D-W)z}{2d^T(z+1)} + \tfrac{z^T(D-W)z}{2d^T(z-1)} = \\ &= \tfrac{\left(z^T(D-W)z\right)d^T 1}{1^T dd^T 1 - z^T d^T d^T z} = \tfrac{\left(z^T(D-W)z\right)d^T 1}{z^T((1^T d)D - dd^T)z} \end{aligned} \tag{4}$$

In the last inequality we used the fact that $1^T d = z^T D z$. When we include general linear constraints on z on the form $Cz = b$, $C \in \mathbb{R}^{m \times n}$, the optimization problem associated with this partitioning cost becomes

$$\begin{aligned} \inf_z \quad & \tfrac{z^T(D-W)z}{z^T((1^T d)D - dd^T)z} \\ \text{s.t.} \quad & z \in \{-1,1\}^n \\ & Cz = b. \end{aligned} \tag{5}$$

The above problem is a non-convex, NP-hard optimization problem. Therefore we are led to replace the $z \in \{-1,1\}^n$ constraint with the norm constraint $z^T z = n$. This gives us the relaxed problem

$$\begin{aligned} \inf_z \quad & \tfrac{z^T(D-W)z}{z^T((1^T d)D - dd^T)z} \\ \text{s.t.} \quad & z^T z = n \\ & Cz = b. \end{aligned} \tag{6}$$

This is also a non-convex problem, however we shall see in section 3 that we are able to solve this problem exactly. Next we will write problem (6) in homogenized form, the reason for doing this will become clear later on. Let L and M be the $(n+1) \times (n+1)$ matrices

$$L = \begin{bmatrix} (D-W) & 0 \\ 0 & 0 \end{bmatrix}, \quad M = \begin{bmatrix} ((1^T d)D - dd^T) & 0 \\ 0 & 0 \end{bmatrix}, \tag{7}$$

and

$$\hat{C} = [C \quad -b] \tag{8}$$

the homogenized constraint matrix. The relaxed problem (6) can now be written

$$\begin{aligned} \inf_z \quad & \tfrac{[z^T \; 1] L \left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right]}{[z^T \; 1] M \left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right]} \\ \text{s.t.} \quad & z^T z = n \\ & \hat{C}\left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right] = 0. \end{aligned} \tag{9}$$

Finally we add the artificial variable z_{n+1}. Let $\hat{z}$ be the extended vector $\left[z^T \;\; z_{n+1}\right]^T$. Throughout the paper we will write $\hat{z}$ when we consider the extended variables and just z when we consider the original variables. The relaxed problem (6) in its homogenized form is

$$\begin{aligned} \inf_{\hat{z}} \quad & \tfrac{\hat{z}^T L \hat{z}}{\hat{z}^T M \hat{z}} \\ \text{s.t.} \quad & \hat{z}_{n+1}^2 - 1 = 0 \\ & \hat{z}^T \hat{z} = n+1 \\ & \hat{C}\hat{z} = 0. \end{aligned} \tag{10}$$

Note that the first constraint is equivalent to $\hat{z}_{n+1} = 1$. If $\hat{z}_{n+1} = -1$ then we may change the sign of $\hat{z}$ to obtain a solution to our original problem.

The homogenized constraints $\hat{C}\hat{z} = 0$ now form a linear subspace and can be eliminated in the following way. Let $N_{\hat{C}}$ be a matrix where its columns form a base of the nullspace of $\hat{C}$. Let $k+1$ be the dimension of the nullspace. Any $\hat{z}$ fulfilling $\hat{C}\hat{z} = 0$ can be written $\hat{z} = N_{\hat{C}}\hat{y}$, where $\hat{y} \in \mathbb{R}^{k+1}$. As in the case with the z-variables, $\hat{y}$ is the vector containing all variables whereas y is a vector containing all but the last variable. Assuming that the linear constraints are feasible we may always choose that basis such that $\hat{y}_{k+1} = \hat{z}_{n+1} = 1$. We put $L_{\hat{C}} = N_{\hat{C}}^T L N_{\hat{C}}$, $M_{\hat{C}} = N_{\hat{C}}^T M N_{\hat{C}}$. In the new space we get the following formulation

$$\begin{array}{ll} \inf_{\hat{y}} & = \frac{\hat{y}^T L_{\hat{C}} \hat{y}}{\hat{y}^T M_{\hat{C}} \hat{y}} \\ \text{s.t.} & \hat{y}_{k+1}^2 - 1 = 0 \\ & \hat{y}^T N_{\hat{C}}^T N_{\hat{C}} \hat{y} = ||\hat{y}||_{N_{\hat{C}}}^2 = n+1, \end{array} \tag{11}$$

we will use $f(\hat{y})$ to denote the objective function of this problem. A common approach to solving this kind of problem is to simply drop one of the two constraints. This may however result in very poor solutions. We shall see that we can in fact solve this problem exactly without excluding any constraints.

3. Lagrangian Relaxation and Strong Duality

In this section we will show how to solve (6) using Lagrange duality. To do this we start by generalizing a lemma from [7] for trust region problems

Lemma 1. *If there exists a y with $y^T A_3 y + 2b_3^T y + c_3 < 0$, then, assuming the existence of a minima, the primal problem*

$$\inf_y \frac{y^T A_1 y + 2b_1^T y + c_1}{y^T A_2 y + 2b_2^T y + c_2}, \; s.t\; y^T A_3 y + 2b_3^T y + c_3 \leq 0 \tag{12}$$

and the dual problem

$$\sup_{\lambda \geq 0} \inf_y \frac{y^T (A_1 + \lambda A_3) y + (b_1 + \lambda b_3)^T y + c_1 + \lambda c_3}{y^T A_2 y + 2b_2^T y + c_2} \tag{13}$$

has no duality gap.

Proof. The primal problem can be written as

$$\begin{array}{ll} \inf & \gamma_1 \\ \text{s.t} & y^T (A_1 - \gamma_1 A_2) y + 2(b_1 - \gamma_1 b_2)^T y + c_1 - \gamma_1 c_2 \leq 0 \\ & y^T A_3 y + 2b_3^T y + c_3 \leq 0 \end{array} \tag{14}$$

Let $\mathcal{M}(\lambda, \gamma)$ be the matrix

$$\mathcal{M}(\lambda, \gamma) = \begin{bmatrix} A_1 + \lambda A_3 - \gamma A_2 & b_1 + \lambda b_3 - \gamma b_2 \\ (b_1 + \lambda b_3 - \gamma b_2)^T & c_1 + \lambda c_3 - \gamma c_2 \end{bmatrix} \tag{15}$$

The dual problem can be written

$$\begin{array}{ll} \sup_{\lambda \geq 0} \inf_{\gamma_2, y} & \gamma_2 \\ \text{s.t} & \begin{bmatrix} y \\ 1 \end{bmatrix}^T \mathcal{M}(\lambda, \gamma_2) \begin{bmatrix} y \\ 1 \end{bmatrix} \leq 0 \end{array} \tag{16}$$

Since (16) is dual to (14) we have that for their optimal values, $\gamma_2^* \leq \gamma_1^*$ must hold. To prove that there is no duality gap we must show that $\gamma_2^* = \gamma_1^*$. We do this by considering the following problem

$$\begin{array}{ll} \sup_{\gamma_3, \lambda \geq 0} & \gamma_3 \\ \text{s.t} & \mathcal{M}(\lambda, \gamma_3) \succeq 0 \end{array} \tag{17}$$

Here $\mathcal{M}(\lambda, \gamma_3) \succeq 0$ means that $\mathcal{M}(\lambda, \gamma_3)$ is positive semidefinite. We note that if $\mathcal{M}(\lambda, \gamma_3) \succeq 0$ then there is no y fulfilling

$$\begin{bmatrix} y \\ 1 \end{bmatrix}^T \mathcal{M}(\lambda, \gamma_3) \begin{bmatrix} y \\ 1 \end{bmatrix} + \epsilon \leq 0 \tag{18}$$

for any $\epsilon > 0$. Therefore we must have that the optimal values fulfills $\gamma_3^* \leq \gamma_2^* \leq \gamma_1^*$. To complete the proof we show that $\gamma_3^* = \gamma_1^*$. We note that for any $\gamma \leq \gamma_1^*$ we have that

$$\begin{array}{c} y^T A_3 y + 2b_3^T y + c_3 \leq 0 \Rightarrow \\ y^T (A_1 - \gamma A_2) y + 2(b_1 - \gamma b_2)^T y + c_1 - \gamma c_2 \geq 0 \end{array} \tag{19}$$

However according to the S-procedure [1] this is true if and only if there exists $\lambda \geq 0$ such that $\mathcal{M}(\lambda, \gamma) \succeq 0$. Therefore (γ, λ) is feasible for problem (17) and thus $\gamma_3 = \gamma_1$. □

We note that for a fixed γ the problem

$$\begin{array}{ll} \inf_y & y^T (A_1 - \gamma A_2) y + 2(b_1 - \gamma b_2)^T y + c_1 - \gamma c_2 \\ \text{s.t.} & y^T A_3 y + 2b_3^T y + c_3 \leq 0 \end{array} \tag{20}$$

only has an interior solution if $A_1 - \gamma A_2$ is positive semidefinite. If A_3 is positive semidefinite then we may subtract $k(y^T A_3 y + 2b_3^T y + c_3)$ $(k > 0)$ from the objective function to obtain boundary solutions. This gives us the following corollary

Corollary 1. *Let A_3 be positive semidefinite. If there exists a y with $y^T A_3 y + 2b_3^T y + c_3 < 0$, then the primal problem*

$$\inf_y \frac{y^T A_1 y + 2b_1^T y + c_1}{y^T A_2 y + 2b_2^T y + c_2}, \; s.t.\; y^T A_3 y + 2b_3^T y + c_3 = 0 \tag{21}$$

and the dual problem

$$\sup_{\lambda} \inf_y \frac{y^T (A_1 + \lambda A_3) y + (b_1 + \lambda b_3)^T y + c_1 + \lambda c_3}{y^T A_2 y + 2b_2^T y + c_2} \tag{22}$$

has no duality gap, (once again assuming that a minima exists for the primal problem).

Next we will show how to solve a problem on a form related to (11). Let

$$\hat{A}_1 = \begin{bmatrix} A_1 & b_1 \\ b_1^T & c_1 \end{bmatrix},\ \hat{A}_2 = \begin{bmatrix} A_2 & b_2 \\ b_2^T & c_2 \end{bmatrix},\ \hat{A}_3 = \begin{bmatrix} A_3 & b_3 \\ b_3^T & c_3 \end{bmatrix}$$

Theorem 1. *Assuming the existence of a minima, if $\hat{A}_3$ is positive definite, then the primal problem*

$$\inf_{y^T A_3 y + 2b_3^T y + c_3 = n+1} \frac{y^T A_1 y + 2b_1^T y + c_1}{y^T A_2 y + 2b_2^T y + c_2} = \\ = \inf_{\substack{\hat{y}^T \hat{A}_3 \hat{y} = n+1 \\ y_{n+1}^2 = 1}} \frac{\hat{y}^T \hat{A}_1 \hat{y}}{\hat{y}^T \hat{A}_2 \hat{y}} \tag{23}$$

and its dual

$$\sup_t \inf_{\hat{y}^T \hat{A}_3 \hat{y} = n+1} \frac{\hat{y}^T \hat{A}_1 \hat{y} + t y_{n+1}^2 - t}{\hat{y}^T \hat{A}_2 \hat{y}} \tag{24}$$

has no duality gap.

Proof. Let γ^* be the optimal value of problem (11). Then

$$\begin{aligned}
\gamma^* &= \inf_{\substack{\hat{y}^T \hat{A}_3 \hat{y} = n+1 \\ y_{n+1}^2 = 1}} \frac{\hat{y}^T \hat{A}_1 \hat{y}}{\hat{y}^T \hat{A}_2 \hat{y}} \\
&= \sup_t \inf_{\substack{\hat{y}^T \hat{A}_3 \hat{y} = n+1 \\ y_{n+1}^2 = 1}} \frac{\hat{y}^T \hat{A}_1 \hat{y} + t y_{n+1}^2 - t}{\hat{y}^T \hat{A}_2 \hat{y}} \\
&\geq \sup_t \inf_{\hat{y}^T \hat{A}_3 \hat{y} = n+1} \frac{\hat{y}^T \hat{A}_1 \hat{y} + t y_{n+1}^2 - t}{\hat{y}^T \hat{A}_2 \hat{y}} \\
&\geq \sup_{t,\lambda} \inf_{\hat{y}} \frac{\hat{y}^T \hat{A}_1 \hat{y} + t y_{n+1}^2 - t + \lambda(\hat{y}^T \hat{A}_3 \hat{y} - (n+1))}{\hat{y}^T \hat{A}_2 \hat{y}} \\
&= \sup_{s,\lambda} \inf_{\hat{y}} \\
&\frac{\hat{y}^T \hat{A}_1 \hat{y} + s y_{n+1}^2 - s + \lambda(y^T A_3 y + y_{n+1} 2b_3^T y + c_3 - (n+1))}{\hat{y}^T \hat{A}_2 \hat{y}} = \\
&= \sup_\lambda \inf_{y_{n+1}^2 = 1} \frac{\hat{y}^T \hat{A}_1 \hat{y} + \lambda(y^T A_3 y + 2b_3^T y + c_3 - (n+1))}{\hat{y}^T \hat{A}_2 \hat{y}} \\
&= \sup_\lambda \inf_y \frac{y^T A_1 y + 2b_1^T y + c_1 + \lambda(y^T A_3 y + 2b_3^T y + c_3 - (n+1))}{y^T A_2 y + 2b_2^T y + c_2} \\
&= \gamma^*
\end{aligned} \tag{25}$$

Where we let $s = t + c_3\lambda$. In the last two equalities corollary 1 was used twice. The third row of the above proof gives us that

$$\begin{aligned}
\mu^* &= \sup_t \inf_{\hat{y}^T \hat{A}_3 \hat{y} = n+1} \frac{\hat{y}^T \hat{A}_1 \hat{y} + t y_{n+1}^2 - t}{\hat{y}^T \hat{A}_2 \hat{y}} = \\
&= \sup_t \inf_{\hat{y}^T \hat{A}_3 \hat{y} = n+1} \frac{\hat{y}^T \hat{A}_1 \hat{y} + t y_{n+1}^2 - t \frac{\hat{y}^T \hat{A}_3 \hat{y}}{n+1}}{\hat{y}^T \hat{A}_2 \hat{y}} = \\
&= \sup_t \inf_{\hat{y}^T \hat{A}_3 \hat{y} = n+1} \frac{\hat{y}^T \left(\hat{A}_1 + t \left(\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right] - \frac{\hat{A}_3}{n+1} \right) \right) \hat{y}}{\hat{y}^T \hat{A}_2 \hat{y}}
\end{aligned} \tag{26}$$

□

Finally, since strong duality holds, we can state the following corollary. [1].

Corollary 2. *If t^* and $\hat{y}^*$ solves* (26)*, then $(\hat{y}^*)^T \hat{N} \hat{y}^* = n+1$ and $y_{k+1}^* = 1$. That is, $\hat{y}^*$ is an optimal feasible solution to* (12)

4. The Dual Problem and Constrained Normalized Cuts

Returning to our relaxed problem (11) we start off by introducing the following lemma.

Lemma 2. *L and M are both $(n+1) \times (n+1)$ positive semidefinite matrices of rank $n-1$, both their nullspaces are spanned by $n_1 = [1 \ldots 1\ 0]^T$ and $n_2 = [0 \ldots 0\ 1]^T$. Consequently, $L_{\hat{C}}$ and $M_{\hat{C}}$ are also positive semidefinite.*

Proof. L is the zero-padded positive semidefinite Laplacian matrix of the affinity matrix W and is hence also positive semidefinite. For M it suffices to show that the matrix $(1^T d)D - dd^T$ is p.s.d.

$$\begin{aligned}
v^T((1^T d)D - dd^T)v &= \textstyle\sum_i d_i \sum_j d_j v_j^2 - (\sum_i d_i v_i)^2 \\
&= \textstyle\sum_{i,j} d_i d_j v_j (v_j - v_i) = \sum_i d_i d_i v_i (v_i - v_i) + \\
&+ \textstyle\sum_{i,j<i} d_i d_j v_j (v_j - v_i) + d_j d_i v_i (v_i - v_j) = \\
&\textstyle\sum_{i,j<i} d_i d_j (v_j - v_i)^2 \geq 0,\ \forall v \in \mathbb{R}^n
\end{aligned} \tag{27}$$

The last inequality comes from $d_i > 0$ for all i which means that $(1^T d)D - dd^T$, and thus also M, are positive semidefinite.

The second statement follows since both $Ln_i = Mn_i = 0$ for $i = 1, 2$.

Next, since

$$\begin{aligned}
&v^T L v \geq 0,\ \forall v \in \mathbb{R}^n \Rightarrow v^T L v \geq 0,\ \forall v \in Null(\hat{C}) \Rightarrow \\
&\Rightarrow w^T {N_{\hat{C}}}^T L {N_{\hat{C}}}^T w \geq 0,\ \forall w \in \mathbb{R}^k \Rightarrow \\
&\Rightarrow w^T L_{\hat{C}} w \geq 0,\ w \in \mathbb{R}^k
\end{aligned}$$

it holds that $L_{\hat{C}} \succeq 0$, and similarly for $M_{\hat{C}}$. □

Assuming that the original problem is feasible then we have that, as $f(\hat{y})$ of problem (23) is the quotient of two positive semidefinite quadratic forms and is therefore $f(\hat{y})$ nonnegative, a minima for the relaxed Normalized Cut problem will exist. Theorem 1 states that strong duality holds for a program on the form (23), if a minima exists. Consequently, we can apply the theory from the previous section directly and solve (11) through its dual formulation. Let

$$E_{\hat{C}} = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right] - \frac{N_{\hat{C}}^T N_{\hat{C}}}{n+1} = N_{\hat{C}}^T \begin{bmatrix} -\frac{I}{n+1} & 0 \\ 0 & 1 \end{bmatrix} N_{\hat{C}} \tag{28}$$

and let $\theta(\hat{y}, t)$ denote the Lagrangian function. The dual problem is then

$$\sup_t \inf_{||\hat{y}||^2_{N_{\hat{C}}}=n+1} \theta(\hat{y}, t) = \frac{\hat{y}^T(L_{\hat{C}} + tE_{\hat{C}})\hat{y}}{\hat{y}^T M_{\hat{C}} \hat{y}} \tag{29}$$

The inner minimization is the well known generalized Rayleigh quotient, for which the minima is given by the algebraically smallest generalized eigenvalue[1] of $(L_{\hat{C}}+tE_{\hat{C}})$ and $M_{\hat{C}}$. Letting $\lambda^G_{\min}(t)$ and $v^G_{\min}(t)$, denote the smallest generalized eigenvalue and corresponding generalized eigenvector of $(L_{\hat{C}} + tE_{\hat{C}})$ and $M_{\hat{C}}$ we can write problem (29) as we can write problem (29) as

$$\sup_t \lambda^G_{\min}(L_{\hat{C}} + tE_{\hat{C}}, M_{\hat{C}}). \tag{30}$$

It can easily be shown that the minimizer of the inner problem of (29), is given by a scaling of the generalized eigenvector, $\hat{y}(t) = (||v^G_{\min}(t)||_{N_{\hat{C}}})v^G_{\min}(t)$. The relaxed Normalized Cut problem can thus be solved by finding the maxima of (30). As the objective function is the point-wise infimum of functions linear in t, it is a concave function, as is expected from dual problems. So solving (30) means maximizing a concave function in one variable t, this can be carried out using standard methods for one-dimensional optimization.

Unfortunately, the task of solving large scale generalized eigenvalue problems can be demanding, especially when the matrices involved are dense, as the case is here. This can however be remedied, by exploiting the unique matrix structure we can rewrite the generalized eigenvalue problem as a standard one. First we note that the generalized eigenvalue problem $Av = \lambda Bv$ is equivalent to the standard eigenvalue problem $B^{-1}Av = \lambda v$, if B is non-singular. Furthermore, in large scale applications it is reasonable to assume that the number of variables $n + 1$ is much greater than the number of constraints m. Then the base for the null space of the homogenized linear constraints $N_{\hat{C}}$ can then be written on the form $N_{\hat{C}} = \left[\begin{smallmatrix} c & c_0 \\ I & \end{smallmatrix}\right]$. Now we can write

$$\begin{aligned}
M_{\hat{C}} &= \left[\begin{smallmatrix} c & c_0 \\ I & \end{smallmatrix}\right]^T \left(\begin{bmatrix} ((1^T d)D - dd^T) & 0 \\ 0 & 0 \end{bmatrix}\right) \left[\begin{smallmatrix} c & c_0 \\ I & \end{smallmatrix}\right] = \\
&= \left\{ \begin{matrix} D := \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \\ d := \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \end{matrix} \right\} = \underbrace{\begin{bmatrix} D_2 & 0 \\ 0 & c_0^T D_1 c_0 + 1 \end{bmatrix}}_{\tilde{D}} + \\
&+ \underbrace{\begin{bmatrix} c^T & cd_1 + d_2 & 0 \\ c_0^T & c_0^T d_1 & 1 \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} D_1 & & \\ & 1 & \\ & & -1 \end{bmatrix}}_{S} \begin{bmatrix} c & c_0 \\ d_1^T c^T + d_2^T & d_1^T c_0 \\ 0 & 1 \end{bmatrix} = \\
&= \tilde{D} + VSV^T
\end{aligned} \tag{31}$$

Hence, $M_{\hat{C}}$ is the sum of a positive definite, diagonal matrix $\tilde{D}$ and a low-rank correction VSV^T. As a direct result of the Woodbury matrix identity [5] we can express the inverse of $M_{\hat{C}}$ as

$$\begin{aligned}
M_{\hat{C}}^{-1} &= (\tilde{D} + VSV^T)^{-1} = \\
&= \tilde{D}^{-1}\left(I - V(S^{-1} + V^T\tilde{D}^{-1}V)^{-1}V\tilde{D}^{-1}\right)
\end{aligned} \tag{32}$$

Despite the potentially immense size of the entering matrices, this inverse can be efficiently computed since $\tilde{D}$ is diagonal and the size of the square matrices S and $(S^{-1}+V^T\tilde{D}^{-1}V)$ are both typically manageable and therefore easily inverted. Our generalized eigenvalue problem then turns into the problem of finding the smallest algebraic eigenvalue of the matrix $M_{\hat{C}}^{-1}L_{\hat{C}}$. The dual problem becomes

$$\begin{aligned}
\sup_t \quad & \lambda_{\min}\big((\tilde{D}^{-1}(I - V(S^{-1} + V^T \quad \tilde{D}^{-1}V)^{-1}V\tilde{D}^{-1}) \\
& N_{\hat{C}}^T(L_{\hat{C}} + tE_{\hat{C}})N_{\hat{C}}\big).
\end{aligned} \tag{33}$$

Not only does this reformulation provide us with the more familiar, standard eigenvalue problem but it will also allow for very efficient computations of multiplications of vectors to this matrix. This is a crucial property, since, even though $M_{\hat{C}}^{-1}(L_{\hat{C}} + tE_{\hat{C}})$ is still dense, it is the product and sum of diagonal ($\tilde{D}^{-1}$, $E_{\hat{C}}$), sparse ($L_{\hat{C}}$, $N_{\hat{C}}$) and low rank matrices (V, S^{-1}). It is a very structured matrix to which iterative eigensolvers can successfully be applied.

In certain case it might however occur that the quadratic form in the denominator is only positive semidefinite and thus singular.

These cases are easily detected and must be treated specially. As we then can not invert $M_{\hat{C}}$ and rewrite the problem as a standard eigenvalue problem we must instead work with generalized eigenvalues, as defined in (30). This is preferably avoided as this is typically a more computationally demanding formulation, especially since the entering matrices are dense. Iterative methods for finding generalized methods for structured matrices such as $L_{\hat{C}} + tE$ and $M_{\hat{C}}$, do however exist [10]. Note, that the absence of linear constraints is such a special instance. However, in that case homogenization is completely unnecessary, and (6) with $Cz = b$ removed, is an standard unconstrained generalized Rayleigh quotient and the solution is given by the generalized eigenvalue $\lambda^T_G(D - W, (1^T d)D - dd^T)$.

Now, if t^* and $\hat{y}^* = (||v^G_{\min}(t^*)||_{N_{\hat{C}}})v^G_{\min}(t^*)$ are the optimizers of (29), and (29), corollary 2 certifies that $(y^*)^T N^T_{\hat{C}} N_{\hat{C}} y^* = n + 1$ and that $\hat{y}^*_{k+1} = 1$. With $\hat{z}^* = \begin{bmatrix} z^* \\ \hat{z}^*_{n+1} \end{bmatrix} = N_{\hat{C}}\hat{y}^*$ and $\hat{z}_{n+1} = \hat{y}_{n+1}$, we have that z^* prior to rounding is the minimizer of (6). Thus we have shown how to, through Lagrangian relaxation, solve the relaxed, linearly constrained Normalized Cut problem exactly.

Finally, the solution to the relaxed problem must be discretized in order to obtain a solution to the original binary

[1] By generalized eigenvalue of two matrices A and B we mean finding a $\lambda = \lambda^G(A, B)$ and v, $||v|| = 1$ such that $Av = \lambda Bv$ has a solution.

problem (5). This is typically carried out by applying some rounding scheme to the solution.

4.1. Multi-Class Constrained Normalized Cuts

Multi-class Normalized Cuts is a generalization of (2) for an arbitrary number of partitions.

$$N_{cut}^k = \sum_{l=1}^{k} \frac{cut(A_l, V)}{assoc(A_l, V)} \tag{34}$$

If one minimizes (34) in an iterative fashion, by, given the current k-way partition, finding a new partition while keeping all but two partitions fixed. This procedure is known as the $\alpha - \beta$-swap when used in graph cuts applications, [3]. The associated subproblem at each iteration then becomes

$$\begin{aligned} \tilde{N}_{cut}^k = \frac{cut(A_i, V)}{assoc(A_i, V)} + \\ + \frac{cut(A_j, V)}{assoc(A_j, V)} + \sum_{l \neq i,j} \frac{cut(A_l, V)}{assoc(A_l, V)} = \\ \frac{cut(A_i, V)}{assoc(A_i, V)} + \frac{cut(A_j, V)}{assoc(A_j, V)} + c, \end{aligned} \tag{35}$$

where pixels not labeled i or j are fixed. Consequently, minimizing the multi-class subproblem can be treated similarly to the bipartition problem. At each iteration we have a problem on the form

$$\begin{aligned} \inf_z \quad & f(z) = \tfrac{z^T(D-W)z}{-z^T dd^T z + (1^T d)^2} \\ \text{s.t.} \quad & z \in \{-1, 1\}^n \\ & Cz = b, \end{aligned} \tag{36}$$

where W, D, C and b will be dependent on the current partition and choice of labels to be kept fixed. These matrices are obtained by removing rows and columns corresponding to pixels not labeled i or j, the linear constraints must also be similarly altered to only involve pixels not currently fixed. Given an initial partition, randomly or otherwise, iterating over the possible choices until convergence ensures a multi-class segmentation that fulfills all constraints. There is however no guarantee that this method will avoid getting trapped in local minima and producing a sub-optimal solution, but during the experimental validation this procedure always produced satisfactory results.

5. Experimental Validation

A number of experiments were conducted to evaluate our proposed formulation but also to illustrate how relevant visual information can be incorporated into the segmentation process through non-homogenous, linear constraints and how this can influence the partitioning.

All images were gray-scale of approximately 100-by-100 pixels in size. The affinity matrix was calculated based on edge information, as described in [6]. The one-dimensional maximization over t was carried out using a golden section search, typically requiring $15 - 20$ eigenvalue calculations. The relaxed solution z was discretized by simply thresholding at 0.

Firstly, we compared our approach with the standard Normalized Cut method, fig. 1. Both approaches produce

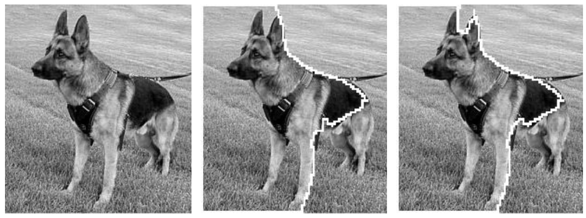

Figure 1. Original image (left), standard Normalized Cut algorithm (middle) and the reformulated Normalized Cut algorithm with no constraints (right).

similar results, suggesting that in the absence of constraints the two formulations are equivalent. However, where our approach has the added advantage of being able to handle linear constraints.

The simplest such constraint might be the hard coding of some pixels, i.e. pixel i should belong to a certain class. This can be expressed as the linear constraints $z_i = \pm 1$, $i = 1..m$. In fig. 2 it can be seen how a number of such hard constraints influences the segmentation of the image in fig. 1.

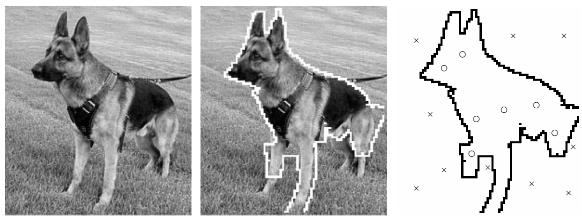

Figure 2. Original image (left), segmentation with constraints (middle) and constraints applied (right).

Another visually significant prior is the size or area of the resulting segments, that is constraints such as $\sum_i z_i = 1^T z = a$. The impact of enforcing limitations on the size of the partitions is shown in fig. 3.

Excluding and including constraints such as, pixel i and j should belong to the same or separate partitions, $z_i + z_j = 0$ or $z_i - z_j = 0$, is yet another meaningful constraint. The result of including a combination of all the above types of constraints can be seen in fig. 4.

Finally, we also performed a multi-class segmentation with linear constraints, fig. 5.

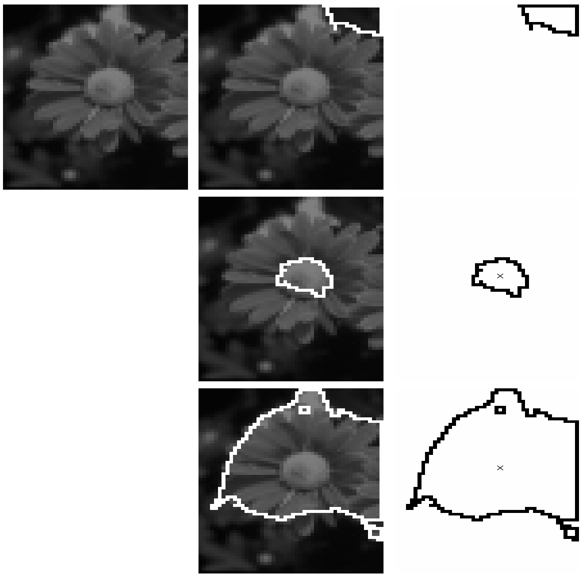

Figure 3. Original image (top left), segmentation without constraints (top middle) and segmentation boundary and constraints applied (top right). Segmentation with area constraints, (area=100 pixels) (middle left), segmentation boundary and constraints applied (middle right). Segmentation with area constraints, (area=2000 pixels) (bottom left), segmentation boundary and constraints applied (bottom right).

We argue that these results, not only indicate a satisfactory performance of the suggested method, but also illustrates the relevance of linear grouping constraints in image segmentation and the impact that they can have on the resulting partitioning. These experiments also seem to indicate that even a simple rounding scheme as the one used here can often suffice. As we threshold at zero, hard, including and excluding constraints are all ensured to hold after discretizing. Only the area constraints are not guaranteed to hold, however probably since the relaxed solution has the correct area, thresholding it typically produces a discrete solution with roughly the correct area.

6. Conclusions

We have presented a reformulation of the classical Normalized Cut problem that allows for the inclusion of linear grouping constraints into the segmentation procedure, through a Lagrangian dual formulation. A method for how to efficiently find such a cut, even for very large scale problems, has also been offered. A number of experiments as well as theoretical proof were also supplied in support of these claims.

Improvements to the presented method include, firstly, the one-dimensional search over t. As the dual function is the point-wise infimum of the eigenvalues of a matrix, it is sub-differentiable and utilizing this information should

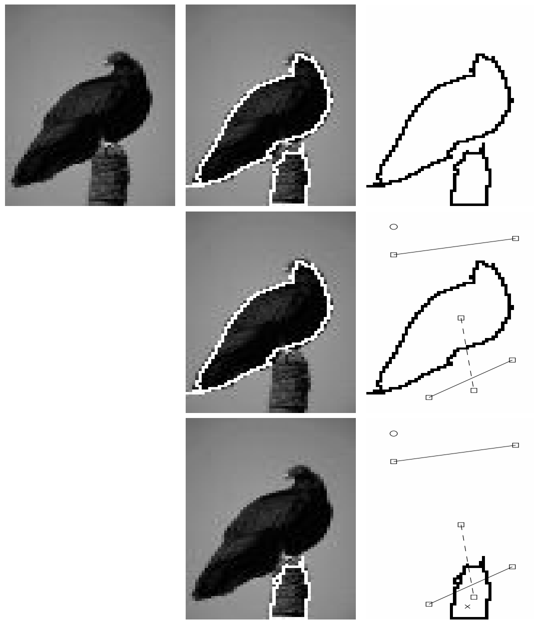

Figure 4. Original image (top left), segmentation without constraints (top middle), segmentation boundary and constraints applied (top right). Segmentation with hard, including and excluding, as well as area constraints, (area=25% of the entire image) (middle left), segmentation boundary and constraints applied (middle right). Segmentation with constraints, (area=250 pixels) (bottom left), segmentation boundary and constraints applied (bottom right). Here a solid line between two pixels indicate an including constraint, and a dashed line an excluding.

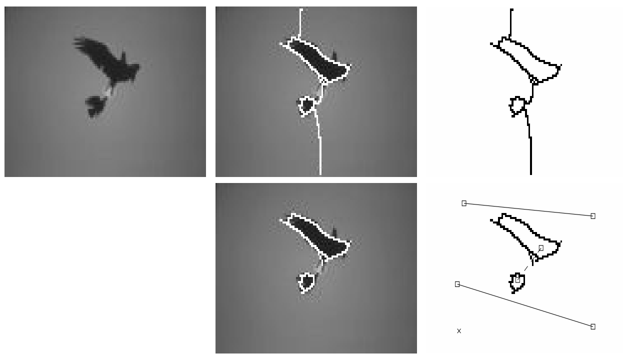

Figure 5. Original image (top left), three-class segmentation without constraints (top middle), segmentation boundary (top right). Three-class segmentation with hard, including and excluding constraints (bottom left), segmentation boundary and constraints applied (bottom right).

greatly reduce the time required for finding t^*. Another issue that was left open in this work is regarding the rounding scheme. The relaxed solution z is currently discretized by simple thresholding at 0. Even though we can guarantee that z prior to rounding fulfills the linear constraints, this is not necessarily true after thresholding and should be addressed. For simpler constraints, as the ones used here, rounding schemes that ensures that the linear constraints hold can easily be devised. We felt that an in-depth discussion on different procedures for discretization was outside the scope of this paper.

Finally, the question of properly initializing the multi-class partitioning should also be investigated as it turns out that this choice can affect both the convergence and the final result.

Acknowledgments

This work has been supported by the European Commission's Sixth Framework Programme under grant no. 011838 as part of the Integrated Project SMErobot^{TM}, Swedish Foundation for Strategic Research (SSF) through the programme Vision in Cognitive Systems II (VISCOS II), Swedish Research Council through grants no. 2004-4579 'Image-Based Localisation and Recognition of Scenes' and no. 2005-3230 'Geometry of multi-camera systems'.

References

[1] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[2] Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. In *International Conference on Computer Vision*, pages 05–112. Vancouver, Canada, 2001.

[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.

[4] T. Cour and J. Shi. Solving markov random fields with spectral relaxation. In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*, volume 11, 2007.

[5] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in Mathematical Sciences, 1996.

[6] 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.

[7] F. Rendl and H. Wolkowicz. A Semidefinite Framework for Trust Region Subproblems with Applications to Large Scale Minimization. Technical Report CORR 94-32, Departement of Combinatorics and Optimization, December 1994.

[8] C. Rother, V. Kolmogorov, and A. Blake. "GrabCut": interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics*, pages 309–314, 2004.

[9] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[10] D. C. Sorensen and C. Yang. Truncated QZ methods for large scale generalized eigenvalue problems. *SIAM Journal on Matrix Analysis and Applications*, 19(4):1045–1073, 1998.

[11] S. Yu and J. Shi. Multiclass spectral clustering. In *International Conference on Computer Vision*. Nice, France, 2003.

[12] S. Yu and J. Shi. Segmentation given partial grouping constraints. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2(26):173–183, 2004.