Results in Applied Mathematics 8 (2020) 100103

Contents lists available at ScienceDirect

Results in Applied Mathematics

journal homepage: www.elsevier.com/locate/results-in-applied-mathematics

A fast constrained image segmentation algorithm

Iván Ojeda-Ruiz, Young-Ju Lee [*,1]

Department of Mathematics, Texas State University, TX, USA[2]

ARTICLE INFO

Article history:
Received 31 December 2019
Received in revised form 7 March 2020
Accepted 31 March 2020
Available online 18 April 2020

ABSTRACT

Normalized cut or Ncut has been one of the most widely used models for image processing. A constraint can also be included in the framework of Ncut to represent a priori information for an effective image segmentation. This results in the so-called constrained Ncut problem. In this paper we present an observation that the constrained Ncut problem can be formulated as an indefinite system of equations under a mild condition on targeted segmentation results. We then show that the indefinite system can be effectively handled by the Augmented Lagrangian Uzawa iterative method together with a classical Algebraic Multigrid Method. Both mathematically and numerically, we demonstrate that the Augmented Lagrangian Uzawa method achieves a solution in one iteration. We show how the proposed method can be efficiently applied for the newly tested recursive two-way Ncut with constraints, i.e., a new constrained sequential segmentation as well. A number of numerical experiments are presented to confirm our theory and to show the superiority of the proposed method. In particular, numerical experiments show that the speed of our algorithm can be orders of magnitude faster than the previously proposed Projected Power Method, a significant improvement of conventional image segmentation algorithms.

1. Introduction

Image segmentation has been an ongoing topic since the 1990's. Many models have been proposed in order to perform accurate image cuts. The Normalized Cut has been a groundbreaking model that has had a huge impact on the accuracy of image cuts. "The Normalized Cut criterion measures both the total dissimilarity between the different groups as well as the total similarity within the groups" explained Shi and Malik when the normalized cut was first proposed in [1]. However, in general, the Normalized Cut has a major drawback whenever there are sections of an image that require more careful selection of the cut. Shi and Yu proposed a modified Normalized Cut model with information on the image cut given *a priori* [2]. A further development was then made by Eriksson et al. [3], which formulate the Normalized Cut segmentation with linear equality constraints in a unified fashion. The idea of using *a priori* information to perform an image cut can be traced back to Marr [4]. It is demonstrated that the constraint can bring the information necessary to overcome the aforementioned major drawback of the Normalized Cut. This is because using *a priori* information can ensure that some parts of an image that are expected to be in one segment stays in the specified parts of the image when cutting the image.

* Corresponding author.
E-mail addresses: ojeda@txstate.edu (I. Ojeda-Ruiz), yjlee@txstate.edu (Y.-J. Lee).

1 The second and corresponding author is supported in part by ACS PRF, USA # 57552-ND9.

2 http://www.txstate.edu/.

https://doi.org/10.1016/j.rinam.2020.100103

Due to the success of the Normalized Cut model with constraints, there have been a number of attempts to efficiently solve the Normalized Cut model with constraints. However, the attempts seem to be based on its original formulation as an eigenvalue problem with constraints. Gander, Golub, and Matt [5] proposed the first method and there have been other attempts to improve the performance of the solution algorithms. One successful method is considered to be the Projected Power Method by Xu et al. in [6].

In this paper, we propose a reformulation of the standard Normalized Cut with constraints into an indefinite system of equations. We then show that the indefinite system can be effectively handled by the Augmented Lagrangian Uzawa iterative method [7] together with a classical Algebraic Multigrid Method [8]. We note that there are in fact, available multiscale image segmentation algorithms (see [9,10] and references cited therein). However, they adopt the different segmentation models. To the authors' best knowledge, the formulation of the original Normalized Cut with constraints as the indefinite system and the solver for the indefinite system based on the Augmented Lagrangian Uzawa iterative method have not been proposed in literature. We demonstrate that the proposed method achieves a solution in one iteration both mathematically and numerically. We show how the proposed method can be efficiently applied for the recursive two-way Normalized Cut with constraints, i.e., sequential constrained segmentations as well. Numerical experiments show that the speed of our algorithm can be orders of magnitude faster than the previously proposed Projected Power Method for the Normalized Cut with constraints.

Throughout the paper, for the sake of simplicity, the null space and the range of any matrix C will be denoted by $\mathcal{N}(C)$ and $\mathcal{R}(C)$, respectively.

The rest of the paper is organized as follows. Section 2 describes the Normalized Cut with constraints, using the notation introduced by Shi and Malik [1]. In Section 3, the constrained minimization problem is reformulated as an indefinite system of equations. A fast solver for the indefinite system is then proposed. A use of the proposed method for recursive two-way Normalized Cut with constraints and a constrained sequential segmentation is discussed. Section 4 presents numerical experiments that demonstrate the performance of the proposed algorithm. The concluding remark and future direction are stated in Section 5.

2. Preliminaries

In order to partition an image (n pixels), we use a graph cut on $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ is the set of vertices (pixels, $|\mathbf{V}| = n$) and $\mathbf{E}$ is the set of edges (connecting a pair of pixels). In order to introduce the Normalized Cut we introduce the following quantities:

$$\text{cut}(\mathbf{X}, \mathbf{V}) = \sum_{a\in\mathbf{X}, b\in\mathbf{V}\setminus\mathbf{X}} w(a, b), \quad \text{and} \quad \text{assoc}(\mathbf{X}, \mathbf{V}) = \sum_{a\in\mathbf{X}, t\in\mathbf{V}} w(a, t), \quad \forall \mathbf{X} \subset \mathbf{V}. \tag{2.1}$$

where $w(a, b)$ is defined as the weight of the edge between a vertex a in $\mathbf{X}$ and a vertex b in $\mathbf{X}\backslash\mathbf{A}$. Note that $w(a, b)$ is expressed by a nonnegative similarity function between nodes a and b. In our paper, we use the Gaussian-Kernel function given below (2.3) following Shi and Malik [1]. Notice that assoc($\mathbf{X}$, $\mathbf{V}$) is the total connection from nodes in $\mathbf{X}$ to all nodes in the graph.

Performing a graph cut implies we partition the set $\mathbf{V}$ into two disjoint sets $\mathbf{A}$ and $\mathbf{B}$ (where $\mathbf{A}\cup\mathbf{B} = \mathbf{V}$ and $\mathbf{A}\cap\mathbf{B} = \emptyset$). The Normalized Cut of $\mathbf{A}$ and $\mathbf{B}$ is then defined as

$$\text{Ncut}(\mathbf{A}, \mathbf{B}) = \frac{\text{cut}(\mathbf{A}, \mathbf{V})}{\text{assoc}(\mathbf{A}, \mathbf{V})} + \frac{\text{cut}(\mathbf{B}, \mathbf{V})}{\text{assoc}(\mathbf{B}, \mathbf{V})}. \tag{2.2}$$

By providing an appropriate labeling of each vertex, we shall identify the label and the vertex. With this identification being introduced, we shall simply write $w(a, b) = w(i, j)$ or simply $w(a, b) = w_{ij}$, where i and j are the label of vertex a and b, respectively. Specifically, the entries of the affinity matrix, $W = (w_{ij})_{i,j=1,\ldots,n}$ are generated with the Gaussian Kernel [1] as follows:

$$w_{ij} = \begin{cases} \exp\left(\dfrac{-|f(i)-f(j)|^2}{\sigma_I^2}\right)\exp\left(\dfrac{-\|\mathbf{x}(i)-\mathbf{x}(j)\|^2}{\sigma_X^2}\right) & \text{if } \|\mathbf{x}(i)-\mathbf{x}(j)\| < r \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

where $\|\cdot\|$ is the Euclidean ℓ^2 norm, r is the parameter, the upper bound of the distance between i and j for which $w_{ij} \neq 0$, $f(i)$ is the brightness value of the ith pixel, $\mathbf{x}(i)$ is the spacial location or coordinate of the ith pixel, σ_I and σ_X are tuning parameters that control the difference in brightness and the difference in spacial location, respectively. Eq. (2.3) should reflect the likelihood that two pixels (nodes i and j) belong to the same point. If the pixels are too far away (farther than a given radius r), then the value of the edge weight w_{ij} is set to 0. It is expected that a larger tuning parameter r produces a better solution for the minimization of the normalized cut. However, we observe that if $r \geq 4$, then visually, sufficient minimization is obtained. Throughout the paper, the numerical results are obtained with the choice of $r = 5$ if r is not specified explicitly. The weight function given in (2.3) is a standard choice and a slightly modified or different choice does not seem to give much different similar results [11].

We define the following quantities given by Shi and Malik:

$$d_i = \sum_j w_{ij}, \quad \kappa = \frac{\sum_{x_i>0} d_i}{\sum_i d_i}, \quad \text{and} \quad \sigma = \frac{\kappa}{1-\kappa} \tag{2.4}$$

and introduce an indicator vector $x \in \mathbb{R}^n$, where $x_i = 1$ if $i \in \mathbf{A}$ and $x_j = -1$ if $j \in \mathbf{B}$, we obtain the following constrained Normalized Cut problem:

$$\min_{x_i \in \{1,-1\}} \text{Ncut}(\mathbf{A}, \mathbf{B})(x) = \min_{u_i \in \{1,-\sigma\}} \frac{u^\top (D-W)u}{u^\top Du} \quad \text{subject to } u^\top De = 0, \tag{2.5}$$

where $u = (1+x) - \sigma(1-x)$, $e = (1, \ldots, 1)^\top \in \mathbb{R}^n$ and $D = \text{diag}(d_i)$. We note that with $v := D^{\frac{1}{2}}u$, the problem (2.5) can be transformed into the following constrained minimization problem:

$$\min_{v_i \in \{d_i^{1/2}, -d_i^{1/2}\sigma\}} \frac{v^\top D^{-\frac{1}{2}}(D-W)D^{-\frac{1}{2}}v}{v^\top v} \quad \text{subject to } v^\top D^{\frac{1}{2}}e = 0. \tag{2.6}$$

Shi and Malik then relax the constraints that $v_i \in \left\{d_i^{1/2}, -d_i^{1/2}\sigma\right\}$ and simply allow v_i to take on any real values. This leads to an eigenvalue problem given as follows:

$$D^{-1/2}(D-W)D^{-1/2}v = \lambda v.$$

The computation is performed to find the eigenvector that corresponds to the smallest nonzero eigenvalue. Once the eigenvector is found, the signs of each component of the eigenvector are used to partition the graph.

Our approach follows the one from Shi and Malik, however, we notice that if two partitions **A** and **B** have the same number of vertices, then $\sigma = 1$ in (2.4) and therefore, u_i is either 1 or -1. This then indicates that the denominator of the objective function, i.e., $u^\top Du$ behaves like a constant.

Additionally, by providing prior information into the minimization problem, i.e., by adding extra constraints to problem (2.6), we expect to get better segmentation results. Mathematically, the constraints can be imposed by introducing a matrix $B \in \mathbb{R}^{m\times n}$ and a vector $c \in \mathbb{R}^m$. B will consist of m rows with zeros everywhere except at the location of the pixel chosen by the user (in lexicographical order) where the entry is 1. The vector c contains the information of the graph to which that pixel belongs. More precisely, we construct the matrix $B = (B_{ij})_{i=1,\ldots,m, j=1,\ldots,n}$ consisting of entries of either one or zero (each row of B has one and only one nonzero entry) and $c = (c_i)_{i=1,\ldots,m}$ consisting of ones and negative ones. For example, for a given kth row of B, let j_k be the column of B for which $B_{k,j_k} = 1$, for $k = 1, \ldots, m$ and let c_k be chosen to be one. This means, we impose the pixel v_{j_k} belongs to the group **A**. On the other hand, if c_k is chosen to be the negative one, it means v_{j_k} is constrained to belong to the group **B**. It is expected that providing more constraints makes more accurate segmentation result. The constraint that ensures some pixels belong to the partition will be chosen by the user. Getting a better result with providing sufficiently small amount of constraints is ideal. Namely, it is desired that $m \ll n$. This discussion leads to the following rather relaxed, but constrained optimization problem.

$$\min_v v^\top D^{-\frac{1}{2}}(D-W)D^{-\frac{1}{2}}v \quad \text{subject to} \quad v^\top D^{\frac{1}{2}}e = 0, \quad Bv = c. \tag{2.7}$$

We remark that L. Xu et al. [6] used the Projected Power Method for solving the problem (2.7) with additional constraints $\|v\| = 1$, which looks to be more faithful for the original Ncut problem. However, the issue arises here that c has to be carefully modified so that two constraints $Bv = c$ and $\|v\| = 1$ do not lead to an ill-posed problem.

The main contribution of this paper is to design a fast solver for the above optimization problem and to show that such a relaxation does indeed a good segmentation result. Additionally, we extend this idea to segment more than two cuts. Namely, we consider k-way Ncut problem via recursive two-way Ncut with constraints. The proposed method can be considered as a new multi-class constrained Ncut. Furthermore, we show that as the recursion progresses, the size of the minimization problem becomes much more smaller while the structure of the problem is maintained.

3. Algorithms for constrained minimization

In this section, we shall reformulate problem (2.7) as an indefinite system of equations and introduce the augmented Lagrangian Uzawa iterative method. We shall demonstrate with mathematical rigor, that the proposed augmented Lagrangian Uzawa method for the indefinite system can attain the solution in one iteration. We begin our discussion with the reformulation of problem (2.7) as an indefinite system of equations.

3.1. An indefinite system equivalent to problem (2.7)

For a fixed $r > 0$, let $W = (w_{ij})$ be given by (2.3) and set

$$A = D^{-1/2}(D-W)D^{-1/2} \in \mathbb{R}^{n\times n} \tag{3.1}$$

Note that the larger the r is, the denser the matrix A is. For $r = \sqrt{2}$, it corresponds to a 9-point stencil. In general, w_{ij} decreases exponentially when $|\mathbf{x}(i) - \mathbf{x}(j)|$ becomes larger. We note that the matrix A is called, a graph Laplacian. It is symmetric and irreducible, i.e., the graph corresponding to A is connected. We can also establish that

$$v^\top (D - W)v = \sum_{i<j} w_{ij}(v_i - v_j)^2. \tag{3.2}$$

This identity shows that A is semidefinite. However, it is not positive definite. In fact, it holds that

$$\mathcal{N}(D - W) = \text{span}\left\{(1, 1, \ldots, 1)^\top \in \mathbb{R}^n\right\} \quad \text{and} \quad \mathcal{N}(A) = \{D^{1/2}e : e \in \mathcal{N}(D - W)\}. \tag{3.3}$$

To reformulate the Problem (2.7), we introduce $B \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$ (typically, $m \ll n$) and consider a functional $F : \mathbb{R}^n \mapsto \mathbb{R}$ defined as follows:

$$F(v) = \frac{1}{2} v^\top A v, \tag{3.4}$$

Then, Problem (2.7) can be formulated as follows:

$$\min_{v\in\mathcal{B}_c} F(v), \text{ where } \mathcal{B}_c = \{v \in \mathbb{R}^n : Bv = c\}. \tag{3.5}$$

Note that the matrix $B = (B_{ij})_{i=1,\ldots,m,j=1,\ldots,n}$ consists of entries of either one or zero. Each row of B has one and only one nonzero entry. Therefore, it is important to observe that B is onto and with e^ℓ, the standard ℓ^{th} Euclidean basis for $\mathbb{R}^n$. Furthermore, we can characterize the null space of B as follows. Let $\{j_1, j_2, \ldots, j_n\}$ is the permutation of $\{1, 2, \ldots, n\}$ so that $B_{kj_k} = 1$ for $1 \le k \le m$,

$$\mathcal{N}(B) = \text{span}\left\{e^{j_{m+1}}, e^{j_{m+2}}, \ldots, e^{j_n} \in \mathbb{R}^n : B_{kj_\ell} = 0, \text{ for } 1 \le k \le m \text{ and } m + 1 \le \ell \le n\right\}. \tag{3.6}$$

Throughout this paper, we let $\{j_{m+1}, j_{m+2}, \ldots, j_n\}$ be the index set that represents the null space of B.

We now introduce a simple, but important conclusion from both characterization of $\mathcal{N}(A)$ and $\mathcal{N}(B)$. Namely,

$$\mathcal{N}(A) \cap \mathcal{N}(B) = \{0\}. \tag{3.7}$$

We are now in a position to present the main result in this section:

Theorem 3.1. *We assume that $B \in \mathbb{R}^{n\times m}$ is onto and A is semidefinite such that $\mathcal{N}(A)\cap\mathcal{N}(B) = \{0\}$. Then, a unique minimizer for the minimization problem (3.5) can be found by solving the following indefinite problem:*

$$\begin{pmatrix} A & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} v \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \tag{3.8}$$

where $f := -Aw$ with w being a particular vector such that $Bw = c$. Namely, the unique minimizer will be $v + w$.

Proof. Let $w \in \mathcal{B}_c$ be given, i.e., $Bw = c$. Then, it is easy to see that the space $\mathcal{B}_c$ is the affine space given by

$$\mathcal{B}_c = w + \mathcal{B}_0, \tag{3.9}$$

and

$$\mathcal{B}_0 = \{v \in \mathbb{R}^n : Bv = 0\}. \tag{3.10}$$

Note that the space $\mathcal{B}_0$ is identical to $\mathcal{N}(B)$. Using this relation, the minimization problem (3.5) can be stated as follows:

$$\min_{v\in\mathcal{B}_0} F(v + w) \quad \text{or} \quad \min_{v\in\mathcal{B}_0} F(v) + v^\top A w + \frac{1}{2} w^\top A w. \tag{3.11}$$

On the other hand, since $w^\top Aw$ is independent of v, the minimization problem (3.11) is equivalent to

$$\min_{v\in\mathcal{B}_0} \frac{1}{2} v^\top A v - v^\top f, \quad \text{where} \quad f := -Aw. \tag{3.12}$$

Although it may not be unique, if v is a minimizer in $\mathcal{B}_0$ for the functional $G(v) = \frac{1}{2}v^\top Av - v^\top f$, then it holds that for any given $u \in \mathcal{B}_0$ and $\varepsilon \in \mathbb{R}$, the function $g(\varepsilon) = G(v + \varepsilon u)$ takes its minimum at $\varepsilon = 0$, i.e.,

$$\frac{dg}{d\varepsilon}(0) = 0 = u^\top (Av - f). \tag{3.13}$$

Due to the fact that u is arbitrary, we conclude that

$$u^\top (Av - f) = 0, \quad \forall u \in \mathcal{N}(B). \tag{3.14}$$

This means $f - Av$ is orthogonal to $\mathcal{N}(B)$. Invoking the closed range theorem stating that

$$\mathcal{N}(B)^\perp = \mathcal{R}(B^\top) \tag{3.15}$$

Table 4.1
Comparison of time elapsed (in seconds (s)) for each algorithm. The parameter μ for the Augmented Lagrangian Uzawa is chosen to be 10^8.

Picture	Projected Power Method	MINRES	Augmented Lagrangian Uzawa
Bird	93.4 s	0.36 s	0.19 s
Camel	398 s	0.68 s	0.30 s
Elephants	601.5 s	19.6 s	3.6 s

Table 4.2
Comparison of number of iterations for each algorithm. The parameter μ for the Augmented Lagrangian Uzawa is chosen to be 10^8.

Picture	Projected Power Method	MINRES	Augmented Lagrangian Uzawa
Bird	9300	206	1
Camel	16,000	238	1
Elephants	22,000	433	1

and using the assumption that B is onto and so $B^\top : \mathbb{R}^m \mapsto \mathbb{R}^n$ is one to one, we arrive at the unique existence of a vector p in $\mathbb{R}^m$ such that

$$B^\top p = f - Av.$$

This means that if we solve the following indefinite problem,

$$f - Av = B^\top p \quad \text{and} \quad Bv = 0, \tag{3.16}$$

then the minimizer for (3.5) is given by $v + w$. It remains to show that in fact, the minimizer v is unique. We assume that there are two minimizers, for instance v_1 and v_2. Then the characterization of the minimizer indicates that

$$v_1 = v_2 + n \quad \text{for some } n \in \mathcal{N}(A). \tag{3.17}$$

But, $Bv_1 = Bv_2 + Bn = Bv_2$. Therefore, $Bn = 0$. On the other hand, $An = 0$ and $\mathcal{N}(A) \cap \mathcal{N}(B) = \{0\}$. Therefore, $n = 0$. This completes the proof. □

Remark 3.1. In practice, we will not consider the case when $\mathcal{N}(A) \cap \mathcal{N}(B) \neq \{0\}$. But, if it is true then the minimizer can be shown to belong to $\mathcal{N}(A)$.

3.2. The augmented Lagrangian Uzawa method for the indefinite system (3.8)

We now discuss how to solve the indefinite system. Let $\mathcal{A} = \begin{pmatrix} A & B^\top \\ B & 0 \end{pmatrix}$. Since A is not invertible, we can introduce the pseudo inverse of A, denoted by $A^\dagger$ [12] and present a simple, but important result.

Lemma 3.1. *The indefinite matrix $\mathcal{A}$ is spectrally equivalent to the mapping*

$$\begin{pmatrix} A & 0 \\ 0 & -BA^\dagger B^\top \end{pmatrix}. \tag{3.18}$$

Proof. We have the following block LU factorization of $\mathcal{A}$:

$$\begin{pmatrix} A & B^\top \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^\dagger & I \end{pmatrix} \begin{pmatrix} A & B^\top \\ 0 & -BA^\dagger B^\top \end{pmatrix}. \tag{3.19}$$

This completes the proof. □

We note that one can try to use the (preconditioned) MINRES to solve the system matrix $\mathcal{A}$ with an appropriate preconditioner. For example, A can be handled by algebraic multigrid method [13]. As demonstrated in Tables 4.1 and 4.2, the performance of MINRES is not too bad in comparison with the Projected Power Method. However, the uniform iteration number cannot be expected for different images without applying appropriate preconditioners. The construction of good preconditioner seems to be, however, unclear yet, for $BA^\dagger B^\top$. Therefore, we take a different path, i.e., Uzawa-type algorithm. In particular, we shall design the augmented Lagrangian Uzawa iterative method. Surprisingly, this method produces the fastest method while the block matrix can be handled by a simple application of algebraic multigrid method.

In passing to the discussion of the Augmented Lagrangian Uzawa iterative method, we point out that the matrix $BA^\dagger B^\top$ is positive definite. For the Ncut with constraints, we see that $BA^\dagger B^\top$ is positive definite. To see this, for any given $x \in \mathbb{R}^m$, we assume that the Euclidean inner product given as follows is zero:

$$0 = (BA^\dagger B^\top x, x) = (A^\dagger B^\top x, B^\top x) = (AA^\dagger A^{-2} AB^\top x, B^\top x) = (A^\dagger A^{-2} AB^\top x, AB^\top x). \tag{3.20}$$

Then, we have that $AB^\top x = 0$. Now, since if $B^\top x = 0$, then $x = 0$, it should hold that $B^\top x = e$. However, $B^\top$ must contains zero rows and therefore, for any $x \in \mathbb{R}^n$, it is impossible to have $B^\top x = e$. This means x must be zero to satisfy Eq. (3.20). This proves that $BA^\dagger B^\top$ is positive definite. Namely, the minimum eigenvalue of $BA^\dagger B^\top$, denoted by μ_0 is positive.

We begin our discussion for Uzawa method applied to the system (3.8). It is well-known that the Uzawa method [14] is based on two steps: (1) solve for u by appropriately inverting the matrix A and apply the Richardson iterative method [15] for the Schur complement system $BA^\dagger B^\top$. The Richardson iterative method converges if we choose damping parameter, say κ so that $0 < \kappa < \frac{2}{\rho(S)}$, where $S = BA^\dagger B^\top$. Certainly, the choice of the damping parameter κ affects the convergence of the Uzawa's iteration and to come up with a procedure for choosing optimal damping parameter is a non-trivial task if at all possible. One remedy is to use the Augmented Lagrangian method [7,16], which solves a modified problem, equivalent to (3.8) by the Uzawa method. The modified problem is as follows:

$$\begin{cases} (A + \mu B^* B)u + B^\top p &= f, \\ Bu &= 0. \end{cases} \tag{3.21}$$

The application of Uzawa for the above system requires two investigations. First, the augmented matrix $A + \mu B^\top B$ is invertible. Second, the spectrum of the Schur complement operator for the indefinite matrix given by

$$\begin{pmatrix} A + \mu B^* B & B^\top \\ B & 0 \end{pmatrix}. \tag{3.22}$$

Typically, we have in mind that the parameter μ will be chosen so that $\mu \gg 1$. As we shall see below, the larger the μ is, the better the convergence speed is. We first discuss the invertibility of the block matrix $A + \mu B^\top B$ for any given $\mu > 0$. In fact, we show that $A + \mu B^\top B$ is M-matrix [17]. We recall that a matrix C is an M-matrix if it is irreducible, i.e., the graph corresponding to C is connected and the following conditions hold:

$$C_{jj} > 0, \quad \forall 1 \le j \le n \tag{3.23a}$$

$$C_{ij} \le 0, \quad \forall i, j : i \ne j \tag{3.23b}$$

$$C_{jj} \ge \sum_{i=1:\neq j}^{n} |C_{ij}|, \quad \forall 1 \le j \le n \tag{3.23c}$$

$$C_{jj} > \sum_{i=1:\neq j}^{n} |C_{ij}|, \quad \text{for at least one } j. \tag{3.23d}$$

Lemma 3.2. *The matrix $A + \mu B^\top B$ is an M-matrix.*

Proof. The matrix A is a graph Laplacian [18,19]. Therefore, it is irreducible and semidefinite. From the property of B, it is clear that $B^\top B$ is the diagonal matrix such that $\mathcal{N}(B^\top B) = \mathcal{N}(B)$. This means that nonzero entries in the diagonal of $B^\top B$ are located in i such that $1 \le i \le n$ and $i \in \{j_1, j_2, \ldots, j_m\}$. Therefore, if $\mu > 0$, then clearly, the matrix $A + \mu B^\top B$ satisfies all the condition that M-matrix should satisfy. This completes the proof. □

Due to Lemma 3.2, we can write the Schur compliment matrix as follows:

$$S_\mu = B(A + \mu B^\top B)^{-1} B^\top. \tag{3.24}$$

Therefore, the application of Sherman–Morrison–Woodbury formula (with a slight modification due that A is not invertible), leads to the following identity:

$$S_\mu^{-1} = [B(A + \mu B^\top B)^{-1} B^\top]^{-1} = \mu I + (BA^\dagger B^\top)^{-1} \tag{3.25}$$

and so, we have that the spectrum of S_μ denoted by $\sigma(S_\mu)$ is given by $\sigma(S_\mu) = \{\eta_0/(1 + \mu\eta_0) : \eta_0 \in \sigma(BA^\dagger B^\top)\}$. Therefore, the spectral radius of S_μ has the upper bound, i.e., $\rho(S_\mu) < 1/\mu$ for any $\mu > 0$. The application of the Uzawa method to (3.21) with damping parameter $\kappa = \mu$ reads: Given (u^ℓ, p^ℓ), the new iterate $(u^{\ell+1}, p^{\ell+1})$ is obtained by solving the following system:

$$\begin{cases} (A + \mu B^\top B)u^{\ell+1} &= f - B^\top p^\ell, \\ p^{\ell+1} &= p^\ell + \mu B u^{\ell+1}. \end{cases} \tag{3.26}$$

Convergence of this method has been discussed in several works [7,16]. Here we also provide a result on the rate of convergence of the Augmented Lagrangian method with references, which indicates that for large μ, the iterates converge very fast to the solution of (3.21). The only difference of the following result from the standard results in [7,16] is that A is not invertible.

Lemma 3.3. *Let (u^0, p^0) be a given initial guess and for $\ell \ge 1$, let (u^ℓ, p^ℓ) be the iterates obtained via the Augmented Lagrangian method. Then the following estimates hold [7,16]:*

$$\|p - p^\ell\| \le \left(\frac{1}{1+\mu\mu_0}\right)^\ell \|p - p^0\|,$$

$$|u - u^\ell|_A \le \sqrt{1/\mu}\|p - p^{\ell-1}\| \le \sqrt{1/\mu}\left(\frac{1}{1+\mu\mu_0}\right)^\ell \|p - p^0\|.$$

where μ_0 is the minimum eigenvalue of $S = BA^\dagger B^\top$ and $|\cdot|_A$ is the A-seminorm.

Remark 3.2. Shown is that $\{u^\ell\}$ converges to u in A-semi-norm. However, it can be converted into a norm. Note that the original system (3.8) is equivalent to a system A replaced by $A + \delta B^*B$ for any $\delta > 0$ since $Bu = 0$. One can then apply the Augmented Lagrangian Uzawa method for the modified system. Then, the A semi-norm will be replaced by $A + \delta B^*B$ norm.

The payoff, however, is that to obtain $u^{\ell+1}$ we need to solve a nearly singular system, such as

$$(A + \mu B^\top B)u^{\ell+1} = f - \mu B^\top p^\ell. \tag{3.27}$$

While the dependency of the number of iterations can be handled by an appropriate multigrid decomposition of the space $\mathbb{R}^n$, in many applications, the key issue arises from the fact that $B^\top B$ would be a *singular* matrix with large null spaces.

Let the algebraic multigrid method generate the nested multigrid structure for the space $\mathbb{R}^n$, say $V_1 \subset V_2 \subset \cdots V_{J-1} \subset V_J$ with $V_J = \mathbb{R}^n$ and each space V_k is further decomposed as a local space so that smoothing can be done in the chosen local spaces, V_ν^ℓ with $\ell = 1, \ldots, n_\nu$. Therefore, the complete space decomposition of $\mathbb{R}^n$ can be given as follows:

$$\mathbb{R}^n = \sum_{\nu=1}^{J} V_\nu = \sum_{\nu=1}^{J}\sum_{\ell=1}^{n_\nu} V_\nu^\ell. \tag{3.28}$$

Generally, the dependency of the performance of standard iterative method is hampered by the large null space of B, or the parameter μ and the problem size, i.e., n. The well-constructed algebraic multigrid method can handle the issue of the size, n using the nested space decomposition. The key issue is in the dependency of the performance on μ. Theory developed in [16,20] indicates that it can be effectively handled by an appropriate local subspace decomposition, i.e.,

$$\mathcal{N}(B) = \sum_{\nu=1}^{J}\sum_{\ell=1}^{n_\nu} V_\nu^\ell \cap \mathcal{N}(B). \tag{3.29}$$

As described in (3.6), $\mathcal{N}(B) = \sum_{k=m+1}^{n} \text{span}\{e^{j_k}\}$. This means if for example, V_J is decomposed so that with $n_J = n$,

$$\mathbb{R}^n = V_J = \sum_{\ell=1}^{n_J} V_J^\ell \quad \text{with } V_J^\ell = \text{span}\{e^\ell\}.$$

With this decomposition of V_J, we see that both identities (3.28) and (3.29) hold true. The algebraic multigrid method based on the above decomposition with exact or inexact local solves for each local subspaces V_ν^ℓ (the former is the point-Gauss–Seidel smoothing) leads to the method that achieves convergence independent of μ. In fact, this means that any standard algebraic multigrid method can achieve the parameter independent convergence with respect to μ.

Lastly, we note that the standard algebraic multigrid method may suffer from the denseness of the matrix. As we have already discussed, the larger the r is the denser the matrix A is. However, since our numerical experiments indicate that if r is sufficiently large, say $r = 4$ or 5, then numerically the segmentation result does the sufficiently good performance and the performance of the algebraic multigrid method for $r = 5$ is sufficiently good, the improvement of algebraic multigrid method for the dense matrix $A + \mu B^\top B$, was not considered in this paper. This is in fact out of the scope of this paper.

3.3. Constrained sequential segmentation

As a simple extension of two-way Ncut with constraints, we can consider constrained sequential segmentations for multiple Ncut problems. In this section, we discuss how our algorithm proposed in Section 3.2 can be applied for the multiple Ncut problem in the framework of constrained sequential segmentation. Let an image be given as well as its graph representation, $\mathbf{G}_0 = (\mathbf{V}_0, \mathbf{E}_0)$. We assume that the following p decompositions of $\mathbf{V}_0$ are considered:

$$\mathbf{V}_0 = \cup_{\ell=1}^{p}\mathbf{A}_\ell \text{ such that } \mathbf{A}_i \cap \mathbf{A}_j = \emptyset \text{ for } i \ne j. \tag{3.30}$$

The algorithm of the recursive two-way Ncut for the p decomposition problem can be described as follows:

Algorithm 3.1. Set $\mathbf{A}_0 = \emptyset$. We then perform:
For $k = 1, \cdots, p$,

(1) Set $\mathbf{A} = \mathbf{A}_k$ and $\mathbf{B} = \cup_{\ell=k+1}^{p}\mathbf{A}_\ell$

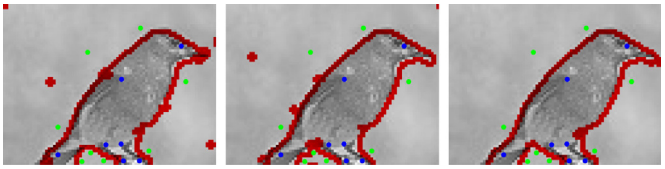

Fig. 4.1. Bird segmentation. Image segmented for $r = 1, r = 2$ and $r = 4$.

(2) Construct $\mathbf{G} = \mathbf{G}_k = (\mathbf{V}_k, \mathbf{E}_k)$ by removing $\mathbf{A}_{k-1}$ as follows:

$$\mathbf{G} = (\mathbf{V}, \mathbf{E}), \text{ where } \mathbf{V} = \mathbf{V}_{k-1} \backslash \mathbf{A}_{k-1} \text{ and } \mathbf{E} = \{\mathbf{e}_{ab} \in \mathbf{E}_{k-1} : \forall a, b \in \mathbf{V}_k\}. \tag{3.31}$$

(3) Construct W and D from $\mathbf{G}$ as well as the constraint for the two Ncut segmentation for $\mathbf{A}$ and $\mathbf{B}$, i.e., the matrix B and the vector c.

(4) Solve the following constrained two-way Ncut problem:

$$\min_{x_i \in \{1,-1\}} \text{Ncut}(\mathbf{A}, \mathbf{B})(x) \text{ subject to } Bv = c. \tag{3.32}$$

EndFor

We would like to remark that the constrained minimization problem in (3.32) can be performed using the proposed algorithm in Section 3.2. In fact, as the recursion becomes deeper, the problem size becomes significantly smaller. Therefore, it is easier to obtain the desired segmentation. Furthermore, the structure of the problem can be maintained in the recursive procedure, which is described below in more details.

The $k-1^{\text{th}}$ step of the Ncut problem is completed with W and D constructed from $\mathbf{G}_{k-1}$. Namely, $\mathbf{A} = \mathbf{A}_{k-1}$ and $\mathbf{B} = \cup_{\ell=k}^{p} \mathbf{A}_\ell$ is assumed to be decomposed. To move on to kth step we generate the permutation matrix $P \in \mathbb{R}^{n_{k-1} \times n_{k-1}}$, where $n_{k-1} = |\mathbf{A} \cup \mathbf{B}|$. More precisely, let $n_k = |\mathbf{B}|$, then the action of P would be given as follows: $Pv = \widehat{v}$, where

$$v = (v_1, v_2, \cdot, v_{n_{k-1}})^\top \text{ and } \widehat{v} = (v_{\sigma_1}, v_{\sigma_2}, \ldots, v_{\sigma_{n_{k-1}}})^\top \text{ with } \sigma_k \in \mathbf{B}, \quad \forall 1 \le k \le n_k. \tag{3.33}$$

In order to generate a new system, we take W obtained from $\mathbf{G}_{k-1}$ and set all entries in σ_k^{th} columns and rows of W for $n_k + 1 \le k \le n_{k-1}$. Using this updated W, denoted by $\overline{W}$, we construct $\overline{D}$ and $M = \overline{D}^{-1/2}(\overline{D} - \overline{W})\overline{D}^{-1/2}$. We then obtain $A \in \mathbb{R}^{n_k \times n_k}$ by taking the first $n_k \times n_k$ block of the matrix $PMP^\top$. Lastly, by constructing the constraint matrix B and a vector c, we arrive at the following constrained minimization:

$$\min_v v^\top A v, \text{ subject to } Bv = c. \tag{3.34}$$

Since the permutation P does not change the structure of the system matrix, the same algorithm in Section 3.2 can be applied.

4. Numerical experiments

In this section, we present the performance of the proposed algorithms, the Augmented Lagrangian Uzawa in comparisons with the standard Projected Power Method and MINRES and also present the result of constrained sequential segmentation.

We first present a simple result that indicates that the choice of r may not be too large. The Fig. 4.1 shows the effect of r by comparing the result for the solutions obtained by using different r's, i.e., $r = 1$, $r = 2$, and $r = 4$. The result indicates while for the small r, the result is not satisfactory, but for relatively large $r \ge 4$, the result immediately improves and stabilizes. Note that the green and blue points are chosen by the user as constraints in order to segment the image. The user is able to choose which points they think are inside of the image and vice versa. The solution contains all of the pixels classified in two graphs. The red line is where the change occurs.

4.1. Performance of iterative methods

We report the performance of the Augmented Lagrangian Uzawa method and MINRES in comparisons with the Projected Power Method [6]. We used a tolerance of 5×10^{-8} for the relative residual error and $r = 5$ for stopping criterion. The CPU times were compared for the total of three pictures, a bird, a camel and a pair of elephants. The size of each picture are 81×64 for the bird, 100×88 for the camel, and 384×256 for the elephant. The results of the CPU time elapsed comparison is shown in Table 4.1.

We also compared the number of iterations for each picture. The results are shown in Table 4.2. As the theory predicts, the number of iteration for the Augmented Lagrangian Uzawa is uniform independent of the images. The chosen parameter for $\mu = 10^8$.

Table 4.3
Iteration number of algebraic multigrid method to solve $A + \mu B^\top B$ for different μ.

Picture	$\mu = 10^2$	$\mu = 10^4$	$\mu = 10^8$
Bird	22 (0.12 s)	21 (0.13 s)	18 (0.18 s)
Camel	21 (0.20 s)	21 (0.20 s)	21 (0.27 s)
Elephants	35 (3.07 s)	35 (3.34 s)	35 (3.58 s)

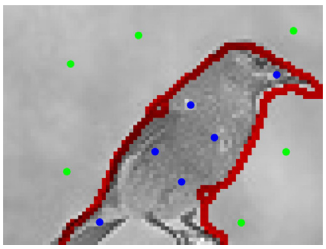

Fig. 4.2. First step constrained segmentation of the bird with points chosen as above. The green and blue points are chosen by the user as constraints in order to segment the image. The red line represents the segmentation produced by the solution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

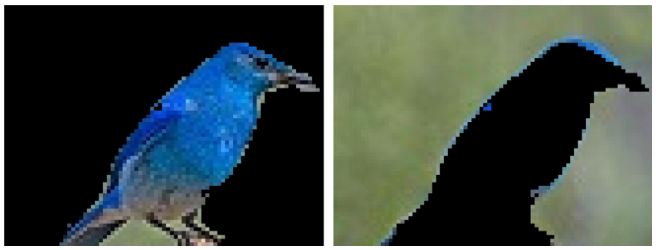

Fig. 4.3. First step constrained segmentation of the bird with points chosen as in 4.2.

Lastly, we present the performance of algebraic multigrid method for solving $A+\mu B^\top B$ for different μ. The performance of the algebraic multigrid method does not deteriorate at all for larger μ, but works robustly as discussed in Section 3.2 (see Table 4.3).

4.2. Results of constrained sequential segmentations with constraints

In this subsection, we present constrained sequential segmentation results. We present the results for the bird and elephant pictures. The bird image has a complication in the bottom where there should be a segmentation between the legs and the tail of the bird. The segmentation may occur incorrectly because the user may be not aware of that section in the image at first sight. In case it happens, the user can apply the second segmentation to fix the problem in a constrained sequential segmentation. This has been described in the following example.

Figs. 4.2 and 4.3 present the first step segmentation result and constraints used in the segmentation (see Fig. 4.2). As discussed, the bottom parts of the bird, between the tail and legs of the bird has not been segmented correctly.

In our second step, we use the left side of Fig. 4.3 and attempt another step of segmentation. New constraints and results are described in Fig. 4.4, which removes the incorrectly segmented part (see Fig. 4.5).

We now consider a bit more challenging example of a pair of elephants. Our goal is to segment large elephant and then small elephant sequentially. The Fig. 4.6 shows the segmentation of the large elephant and points chosen for the segmentation. The left side of Fig. 4.7 is then used for the second step of segmentation, the segmentation of the small elephant. The Fig. 4.8 shows the second step segmentation results together with the points chosen as constraints. This has been post processed and presented in Fig. 4.9. Visually, we successfully achieved the desired segmentation results.

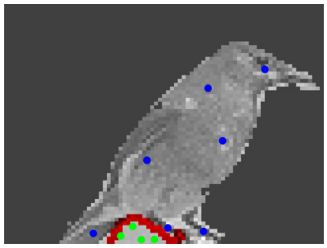

Fig. 4.4. Second step constrained segmentation of the bird with points chosen as above.

Fig. 4.5. Second step constrained segmentation of the bird with points chosen as in 4.4.

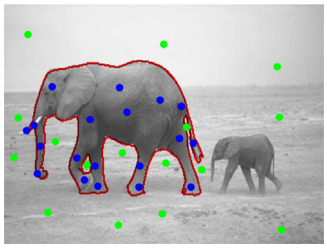

Fig. 4.6. First step constrained segmentation of Elephants with points chosen as above. The green and blue points are chosen by the user as constraints in order to segment the image. The red line represents the segmentation produced by the solution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

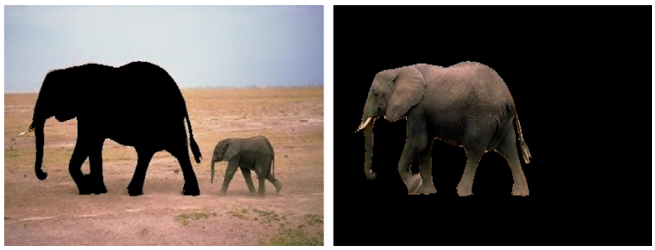

Fig. 4.7. First step constrained segmentation of Elephants with points chosen as chosen in 4.6.

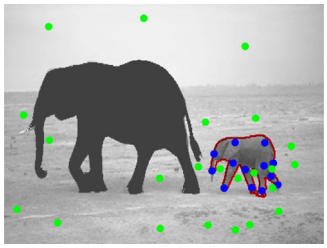

Fig. 4.8. Second step constrained segmentation of Elephants with points chosen as above.

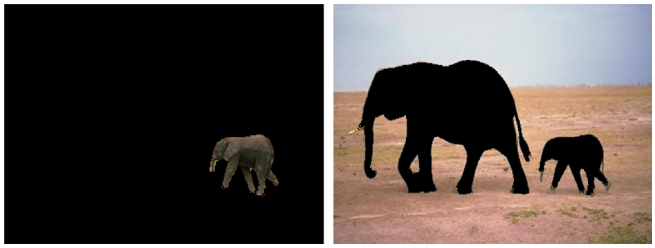

Fig. 4.9. Second step constrained segmentation of Elephants with points chosen as in 4.8.

5. Conclusion

In this paper we have demonstrated that the Normalized Cut problem with linear constraints can be formulated as an indefinite system of equation. We then propose the Augmented Lagrangian Uzawa method to solve the indefinite system. The proposed method shows a significant improvement in computing speed and it allows for the information a priori to be easily implemented. We note that the Augmented Lagrangian Uzawa is a method that can be mathematically proved with uniform convergence and demonstrate that it can be used even for multiple constrained segmentations in a recursive fashion with an advantage of reducing the degree of freedom in each recursion step. In numerical tests, we showed how r can affect the solution to the constrained segmentation problem as well and concluded that a relatively large values of r lead to robust and sufficiently good results. In our future works, we shall design a preconditioner for the MINRES and attempt to understand the relation between multiple Normalized Cut with constraints and the proposed recursive two-way Normalized Cut with constraints.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Authors appreciate the constructive comments from anonymous referees. These comments were crucial to improve the manuscripts much better than the originally submitted manuscript.

References

[1] Shi J, Malik J. Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 2000;22(8):888–905.
[2] Yu S, Shi J. Segmentation given partial grouping constraints. IEEE Trans Pattern Anal Mach Intell 2003;26:173–83.
[3] Eriksson A, Olsson C, Kahl F. Normalized cuts revisited: A reformulation for segmentation with linear grouping constraints. J Math Imaging Vision 2011;39:45–61.
[4] Marr David. Vision. CA: Freeman; 1982.
[5] Gander W, Golub GH, Matt U. A constrained eigenvalue problem. Linear Algebra Appl 1989;322(10):815–39.
[6] Xu L, Li W, Schuurmans D. Fast normalized cut with linear constraints, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 2866–2873.
[7] Fortin Michel, Glowinski Roland. Augmented Lagrangian methods: Applications to the numerical solution of boundary-value problems, Vol. 15. Elsevier; 2000.
[8] Brandt Achi. Algebraic multigrid theory: The symmetric case. Appl Math Comput 1986;19(1–4):23–56.
[9] Sharon E, Galvun M, Sharon D, Basri R, Brandt A. Hierarchy and adaptivity in segmenting visual scenes. Nature 2006;442:810–3.
[10] Sharon E, Brandt A, Basri R. Fast multiscale image segmentation, in: Proc. IEEE Conference on Compututer Vision and Pattern Recognition, vol. 1, 2000, pp. 70–77.
[11] Malik J, Perona P. Preattentive texture discrimination with early vision mechanisms. J Opt Soc Amer A 1990;7:923–32.
[12] Ben-Israel Adi, Greville Thomas NE. Generalized inverses: Theory and applications, Vol. 15. Springer Science & Business Media; 2003.
[13] Xu Jinchao, Zikatanov Ludmil. Algebraic multigrid methods. Acta Numer 2017;26:591–721.
[14] Bramble James H, Pasciak Joseph E, Vassilev Apostol T. Analysis of the inexact uzawa algorithm for saddle point problems. SIAM J Numer Anal 1997;34(3):1072–92.
[15] Saad Y. Numerical methods for large eigenvalue problems. second ed.. 3600 Market Street, 6th Floor, Philadelphia, PA 19104-2688 USA: Society for Industrial and Applied Mathematics; 2011.
[16] Lee Young-Ju, Wu Jinbiao, Xu Jinchao, Zikatanov Ludmil. Robust subspace correction methods for nearly singular systems. Math Models Methods Appl Sci 2007;17(11):1937–63.
[17] Xu Jinchao, Zikatanov Ludmil. A monotone finite element scheme for convection-diffusion equations. Math Comput Am Math Soc 1999;68(228):1429–46.
[18] Pothen A, Simon HD, K.-P Liou. Partitioning sparse matrices with eigenvectors of graphs. SIAM J Matrix Anal Appl 1990;11:430–52.
[19] Biggs N. Algebraic graph theory. Cambridge University Press; 1974, p. 23–7.
[20] Schöberl Joachim. Robust multigrid preconditioning for parameter-dependent problems i: the Stokes-type case. In: Multigrid methods V. Springer; 1998, p. 260–75.