Semidefinite and Second Order Cone Programming Seminar
Fall 2012
Report:
Semidefinite Programming in Computer Vision

Instructor: Farid Alizadeh
Scribe: Svetlana Soloveva

12/10/2012

1 Overview

This is a short review of papers that consider fundamental problems in computer vision solved with semidefinite programming.

2 Image Segmentation

Let us start with Keuchel, J., Schnörr, C., Schellewald, C., and Cremers, D. *Binary Partitioning, Perceptual Grouping, and Restoration with Semidefinite Programming* [2].

2.1 Problem Formulation

Image is represented by an undirected weighted graph $G = (V, E)$ where vertices $i \in V = \{1, \dots, n\}$ correspond to image features, and edges $(i, j) \in E$ have weights $w(i, j) : V \times V \mapsto R^+$ that reflect pairwise similarity between features. To segment an image, we wish to partition nodes V into two disjoint coherent sets S and $\bar{S} = V \setminus S$. This means that edges within the same set should have high weights, and edges between sets should have relatively low weights which results in a minimum weight cut problem.
Let us denote W weighted adjacency matrix of a graph G:

$$W_{ij} = \begin{cases} w(i,j), & \text{if } (i,j) \in E \\ 0, & \text{otherwise.} \end{cases}$$

Denote L Laplacian or Kirchoff matrix of a graph G:
$\mathsf{L} = diag(W\boldsymbol{e}) - W \Leftrightarrow$

$$L_{ij} = \begin{cases} -w(i,j), & \text{if } (i,j) \in E,\ i \neq j, \\ \sum\limits_{j \in V:(i,j) \in E} w(i,j), & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Then weight of a cut δS defined by partition $(S, \bar{S})$ is

$$w(\delta S) = \sum_{i \in S, j \in \bar{S}} w_(i,j) = \frac{1}{4}\mathbf{x}^T L\mathbf{x}.$$

We define the following problem for image segmentation:

$$\begin{cases} \min \mathbf{x}^T L\mathbf{x}, \\ \mathbf{c}^T\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \in \{-1, 1\}^n \end{cases} \tag{1}$$

Here $\mathbf{c}^T\mathbf{x} = \mathbf{b}$ is a "balanced cut" constraint. A simple case $\mathbf{c} = \mathbf{e}, \mathbf{b} = \mathbf{0}$ favours an equal number of vertices in each set.

2.2 Semidefinite Relaxation

First the variables are lifted to a higher dimensional space where the original problem is relaxed to semidefinite programming problem. The lifting step has been introduced in a more general setting by Lovász and Schrijver [5].
Let us rewrite problem (1)in the following form:

$$\begin{cases} \min \mathbf{x}^T L\mathbf{x}, \\ (\mathbf{c}^T\mathbf{x})^2 = \mathbf{b}^2, \\ x_i^2 - 1 = 0, i = 1, \dots, n \end{cases}$$

Observe that the objective function of problem (1) can be rewritten as $L \bullet \mathbf{x}^T\mathbf{x}$, and $\mathbf{x}^T\mathbf{x}$ is positive semidefinite matrix with rank one. Therefore, by introducing the matrix variable $X = \mathbf{x}^T\mathbf{x}$, the problem variables are lifted into a higher dimensional space. Then dropping rank one constraint on matrix X will result the following semidefinite relaxation of problem (1):

$$\begin{cases} z_p = \min L \bullet X, \\ \mathbf{c}\mathbf{c}^T \bullet X = \mathbf{b}^2, \\ diag(X) = I \\ X \succeq 0 \end{cases} \tag{2}$$

Optimal solution X^* of problem (2) can be computed using an interior-point method.

2.3 Solution

Based on X^*, we want to find the optimal solution of original problem (1). For this the method proposed by Goemans and Williamson [1] is used. Since X^* is positive semidefinite, we can compute matrix $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n) : X^* = V^T V$ by Cholesky decomposition. Then we choose random vector $\mathbf{r}$ from the unit sphere and set $x_i^* = 1$ if $\mathbf{v}_i^T \mathbf{r} \geq 0$ and $x_i^* = 0$ otherwise. This procedure is repeated for different random vectors $\mathbf{r}$ to compute $\mathbf{x}^*$ which corresponds to the minimum value of the objective function $\mathbf{x}^T L \mathbf{x}$.
Let us give an interpretation for this method. It follows from the constraint $diag(X) = I$ that $\|\mathbf{v}_i\| = 1$, $i = 1, \ldots, n$. Hence, we associate with each x_i a vector $\mathbf{v}_i$ on the unit sphere. Then by selecting random vectors $\mathbf{r}$, we define hyperplanes that partition vectors $\mathbf{v}_i$ into two sets which at the same time results partition x_i.

2.4 Bounds

By [1] and [2], for the case $\mathbf{c} = \mathbf{0}$, the expected value $E[z^*]$ of the objective function $\mathbf{x}^T L \mathbf{x}$ calculated with the randomized-hyperplane technique is bounded by $E[z^*] \leq \alpha z_p + (1 - \alpha) \sum |L_{ij}|$, where $\alpha \approx 0.878$.

3 Related Work

The work [2] was continued by [3] for segmentation into multiple clusters. In a preprocessing step over-segmentation is performed: coherent image patches are found by mean shift procedure, and a graph is constructed using patches with corresponding mean colors as vertices and edges with weights calculated as sums of the weights of edges between patches. Then SDP relaxation method is applied to find binary segmentations.

Later the approach from [2] was extended for image labeling problems involving multiple classes where labeling was assumed to be a Markov Random Field with maximum a posteriori probability. This problem is equivalent to minimizing energy of the following form:

$$E(x) = \sum_i C(x_i) + \sum_{(i,j)} V(x_i, x_j)$$

where data-fitting term measures cost $C(x_i)$ of assignment element i to class x_i, and smoothness term evaluates disagreement cost $V(x_i, x_j)$ of assignment i to x_i and j to x_j.
As in [2], a trace formulation of the above problem was given. Then labeling matrix was lifted to a higher dimensional space, and semidefinite relaxation was derived.

Schellewald and Schnörr [6] proposed subgraph matching using semidefinite programming. Subgraph matching is applied for object recognition to match a model graph that represents an object and a scene graph that represents extracted image features.
To formulate this problem, let us denote with V_K and V_L set of nodes of a model graph and a scene graph. For each pair of vertices $i \in V_K$ and $j \in V_L$ corresponding edges (i, j) have weights $w(i, j)$ defined by distance function measuring similarity of a pair. Vector $x \in \{0,1\}^{KL}$ is such that x_{ji} indicates the node $i \in V_K$ is matched to the node $j \in V_L$. Matrix $A = (A_K^T, A_L^T)^T$ is the incidence matrix of the bipartite graph $(V_K \bigcup V_L, E)$.
Matrix $Q = N_K \bigotimes \bar{N_L} + \bar{N_K} \bigotimes N_L$ represents relational structures in both model graph and scene graph with their adjacency matrices N_K and N_L. Then optimal matching can be found as a solution of the following problem:

$$\begin{cases} \min \boldsymbol{w}^T \boldsymbol{x} + \alpha \boldsymbol{x}^T Q \boldsymbol{x}, \\ A_K \boldsymbol{x} = \boldsymbol{e}_K, \\ A_L x \leq \boldsymbol{e}_L \\ \boldsymbol{x} \in \{0,1\}^{KL}, \alpha \in R^+ \end{cases} \tag{3}$$

Similarly to section 2.2, we can reformulate the objective function:

$$\boldsymbol{w}^T \boldsymbol{x} + \alpha \boldsymbol{x}^T Q \boldsymbol{x} = \begin{pmatrix} 1 & \boldsymbol{x}^T \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2}\boldsymbol{w}^T \\ \frac{1}{2}\boldsymbol{w} & \alpha Q \end{pmatrix} \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} = \mathrm{tr}(\tilde{Q} X)$$

where $\tilde{Q} \in R^{(KL+1)\times(KL+1)}$, $X \in R^{(KL+1)\times(KL+1)}$,

$$\tilde{Q} = \begin{pmatrix} 0 & \frac{1}{2}\boldsymbol{w}^T \\ \frac{1}{2}\boldsymbol{w} & \alpha Q \end{pmatrix}$$

,

$$X = \begin{pmatrix} 1 & \boldsymbol{x}^T \end{pmatrix} \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix}$$

Note that matrix X is positive semidefinite and has rank 1. Dropping rank 1 constraint on X, results the following semidefinite relaxation which corresponds to lifting of the original problem from vector space with dimension KL into the space of positive semidefinite matrices with dimension $(KL+1) \times (KL+1)$:

$$\begin{cases} \min \mathrm{tr}(\tilde{Q} X), \\ \mathrm{tr}(A_i X) = \boldsymbol{c}_i, \\ X \succeq 0 \end{cases} \tag{4}$$

where constraint matrices $A_i \in R^{(KL+1)\times(KL+1)}$. Opimal solution X^* can be computed with an interior point method. Denote by $x_{sol} \in R^{KL}$ vector of diagonal elements of matrix X^* omitting the first one. Vector x_{sol} is the approximation of the matching vector $x \in \{0,1\}^{KL}$. To find the matching, solve the following problem:

$$\begin{cases} \max \boldsymbol{x}_{sol}^T \boldsymbol{x}, \\ A_K \boldsymbol{x} = \boldsymbol{e}_K, \\ A_L x \leq \boldsymbol{e}_L \\ \boldsymbol{x} \in [0,1]^{KL}, \end{cases} \tag{5}$$

Since incidence matrix of a bipartite graph is totally unimodular, A_K, A_L are totally unimodular, therefore, the optimal solution x^* of problem (5) is integral which represents the optimal matching.

P.H.S. Torr [7] considers solving MRF using semidefinite relaxation and continues [8] with second order cone relaxation of MAP-MRF for subgraph matching. [8] shows SOCP relaxations to be computationally more efficient than SDP relaxations and more accurate than LP relaxations.

4 Conclusion

Discussed methods can be applied to a very general class of problems in computer vision, not restricted to special types of energy functions.
However, due to squared number of variables of the semidefinte relaxation, the computational time grows with the number of image elements and the number of classes so that the application of these methods is yet restricted to small problems with few classes (problems with ten thousands of variables cannot be solved). To overcome this drawback, one could perform successive labeling on different scales. In this way, instead of computing solution for one large problem, a sequence of smaller problems is considered.

References

[1] Goemans, M.X., Williamson, D.P. *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming.* Journal of the ACM. **42(6)** (1995) 1115-1145.

[2] Keuchel, J., Schnörr, C., Schellewald, C., Cremers, D. *Binary Partitioning, Perceptual Grouping, and Restoration with Semidefinite Programming.* Pattern Analysis and Machine Intelligence, IEEE, (2003).

[3] Keuchel, J., Heiler, M., Schnörr, C. *Hierarchical image segmentation based on semidefinite programming.* Pattern Recognition, Springer, (2004).

[4] Keuchel, J. *Multiclass image Labeling with emidefinite programming.* Computer Vision, ECCV (2006), Springer.

[5] Lovász, L., Schrijver, A. *Cones of matrices and set-functions and 0-1 optimization.* SIAM J. Optimization. **1(2)**, (1991), 166-190.

[6] Schellewald, C., Schnörr. C. *Probabilistic subgraph matching based on convex relaxation.* Energy Minimization Methods in Computer Vision and Pattern Recognition, (2005).

[7] P.H.S Torr. *Solving Markov random fields using semidefinite programming.* AISTATS, (2003).

[8] Kumar, P., P.H.S Torr, Zisserman, A. *Solving Markov random fields using second order cone programming relaxations.* omputer Vision and Pattern Recognition, IEEE, (2006).