Supplementary Material for "Semi-supervised Learning for Large Scale Image Cosegmentation"

Zhengxiang Wang Rujie Liu
Fujitsu Research & Development Center Co., Ltd, Beijing, China
{wangzhengxiang,rjliu}@cn.fujitsu.com

Abstract

This supplementary material provides the detailed derivation of the three terms in the energy function to the binary quadratic programming (QP) problem, as well as its decomposition to sub-problems. The plots of the cosegmentation accuracy affected by the different values of parameters are also attached, showing that the proposed method is not sensitive to the choice of parameters within a range.

1. Detailed Equation Derivation

We firstly show how the inter-image distance, intra-image distance and balance term of the energy function (Equation M1, M2 and M5 respectively, M stands for the equation label from the main manuscript, to distinguish the equation label in this supplementary material) are converted to form the binary QP problem (Equation M8, M13 and M16 respectively).

1.1. Inter-image distance

The **inter-image distance** is converted from Equation M1 to Equation M8 by direct expansion:

$$\begin{aligned}
E_{inter} &= \sum_{i=1}^{N_u}\sum_{j=1}^{N_s} \parallel H_i \cdot y_i - H_j^{tr} \cdot y_j^{tr} \parallel^2 + \sum_{i=1}^{N_u}\sum_{j=i+1}^{N_u} \parallel H_i \cdot y_i - H_j \cdot y_j \parallel^2 \\
&= \sum_{i=1}^{N_u}\sum_{j=1}^{N_s} [y_i^T \cdot H_i^T \cdot H_i \cdot y_i - 2 \cdot y_i^T \cdot H_i^T \cdot H_j^{tr} \cdot y_j^{tr} + (y_j^{tr})^T \cdot (H_j^{tr})^T \cdot H_j^{tr} \cdot y_j^{tr}] \\
&+ \sum_{i=1}^{N_u}\sum_{j=i+1}^{N_u} (y_i^T \cdot H_i^T \cdot H_i \cdot y_i - 2 \cdot y_i^T \cdot H_i^T \cdot H_j \cdot y_j + y_j^T \cdot H_j^T \cdot H_j \cdot y_j) \\
&= \sum_{i=1}^{N_u} y_i^T \cdot (N_s \cdot H_i^T \cdot H_i) \cdot y_i + \sum_{i=1}^{N_u} y_i^T \cdot (-2H_i^T \cdot \sum_{j=1}^{N_s} H_j^{tr} \cdot y_j^{tr}) + N_u \cdot \sum_{i=1}^{N_s} (y_i^{tr})^T \cdot (H_i^{tr})^T \cdot H_i^{tr} \cdot y_i^{tr} \\
&+ \sum_{i=1}^{N_u} y_i^T \cdot [(N_u - 1) \cdot H_i^T \cdot H_i] \cdot y_i + \sum_{i=1}^{N_u}\sum_{j=i+1}^{N_u} y_i^T \cdot (-2H_i^T \cdot H_j) \cdot y_j \\
&= \sum_{i=1}^{N_u} y_i^T \cdot [(N_s + N_u - 1) \cdot H_i^T \cdot H_i] \cdot y_i + \sum_{i=1}^{N_u} y_i^T \cdot V_i + C + \sum_{i=1}^{N_u}\sum_{j=i+1}^{N_u} y_i^T \cdot M_{ij}^{inter} \cdot y_j \\
&= \sum_{i=1}^{N_u} y_i^T \cdot M_{ii}^{inter} \cdot y_i + \sum_{i=1}^{N_u}\sum_{j=i+1}^{N_u} y_i^T \cdot M_{ij}^{inter} \cdot y_j + \sum_{i=1}^{N_u} y_i^T \cdot V_i + C
\end{aligned} \tag{1}$$

where M_{ii}^{inter}, M_{ij}^{inter}, V_i and C are defined in Equation M9, M10, M11 and M12 respectively.

1.2. Intra-image distance

The **intra-image distance** is converted from Equation M2 to Equation M13 according to the fact that y_i is a binary vector, we have:

$$|y_i(j) - y_i(k)| = [y_i(j) - y_i(k)]^2 \tag{2}$$

Therefore Equation M2 can be reformulated as:

$$\begin{aligned}
E_{intra} &= \sum_{i=1}^{N_u} \sum_{j=1,k=1}^{s_i} W_i(j,k) \cdot [y_i(j) - y_i(k)]^2 \\
&= \sum_{i=1}^{N_u} \sum_{j=1}^{s_i} \sum_{k \in N(j)} [W_i(j,k) \cdot y_i(j)^2 + W_i(j,k) \cdot y_i(k)^2 - 2W_i(j,k) \cdot y_i(j) \cdot y_i(k)] \\
&= \sum_{i=1}^{N_u} \{ \sum_{j=1}^{s_i} y_i(j)^2 \cdot [\sum_{k \in N(j)} W_i(j,k) + W_i(k,j)] - \sum_{j=1}^{s_i} \sum_{k=1}^{s_i} y_i(j) \cdot [W_i(j,k) + W_i(k,j)] \cdot y_i(k) \}
\end{aligned} \tag{3}$$

We can reformulate the above equation in the matrix form by constructing a matrix M_i^{intra} with the definition provided in Equation M14 and M15, so that the above equation can be converted to Equation M13:

$$\begin{aligned}
E_{intra} &= \sum_{i=1}^{N_u} \{ \sum_{j=1}^{s_i} y_i(j)^2 \cdot [\sum_{k \in N(j)} W_i(j,k) + W_i(k,j)] - \sum_{j=1}^{s_i} \sum_{k=1}^{s_i} y_i(j) \cdot [W_i(j,k) + W_i(k,j)] \cdot y_i(k) \} \\
&= \sum_{i=1}^{N_u} y_i^T \cdot M_i^{intra} \cdot y_i
\end{aligned} \tag{4}$$

1.3. Balance term

The **balance term** is converted from Equation M5 to Equation M16 through Taylor expansion, we firstly rewrite Equation M5 as:

$$E_{bal} = \sum_{i=1}^{N_u} [P_i^f \log P_i^f + (1 - P_i^f) \log(1 - P_i^f)] = \sum_{i=1}^{N_u} f(P_i^f) \tag{5}$$

where

$$f(x) = x \log x + (1 - x) \log(1 - x) \tag{6}$$

Since $x = P_i^f \in [0, 1]$ is continually derivative, we can use Taylor expansion to approximate $f(x)$ at $x = \frac{1}{2}$:

$$\begin{aligned}
f(x) &\approx f(\frac{1}{2}) + f'(\frac{1}{2})(x - \frac{1}{2}) + \frac{f''(\frac{1}{2})}{2}(x - \frac{1}{2})^2 \\
&= -1 + 2(x - \frac{1}{2})^2 \\
&= 2x^2 - 2x - \frac{1}{2}
\end{aligned} \tag{7}$$

Therefore

$$\begin{aligned}
E_{bal} &= \sum_{i=1}^{N_u} f(P_i^f) \\
&= \sum_{i=1}^{N_u} [2(P_i^f)^2 - 2P_i^f - \frac{1}{2}] \\
&= \sum_{i=1}^{N_u} (2 \frac{y_i^T \cdot e_i \cdot e_i^T \cdot y_i}{s_i^2} - 2 \frac{y_i^T \cdot e_i}{s_i} - \frac{1}{2})
\end{aligned} \tag{8}$$

1.4. The whole energy function

By summing all these three terms, **the whole energy function** E of Equation M17 can be derived by:

$$
\begin{aligned}
E &= E_{inter} + \lambda_1 \cdot E_{intra} + \lambda_2 \cdot E_{bal} \qquad (9) \\
&= \sum_{i=1}^{N_u} y_i^T \cdot M_{ii}^{inter} \cdot y_i + \sum_{i=1}^{N_u} \sum_{j=i+1}^{N_u} y_i^T \cdot M_{ij}^{inter} \cdot y_j + \sum_{i=1}^{N_u} y_i^T \cdot V_i + C \\
&+ \lambda_1 \sum_{i=1}^{N_u} y_i^T \cdot M_i^{intra} \cdot y_i + \lambda_2 \sum_{i=1}^{N_u} (2\frac{y_i^T \cdot e_i \cdot e_i^T \cdot y_i}{s_i^2} - 2\frac{y_i^T \cdot e_i}{s_i} - \frac{1}{2}) \\
&= \sum_{i=1}^{N_u} y_i^T \cdot (M_{ii}^{inter} + \lambda_1 M_i^{intra} + 2\lambda_2 \frac{e_i \cdot e_i^T}{s_i^2}) \cdot y_i + \sum_{i=1}^{N_u} \sum_{j=i+1}^{N_u} y_i^T \cdot M_{ij}^{inter} \cdot y_j + \sum_{i=1}^{N_u} y_i^T \cdot (V_i - 2\lambda_2 \frac{e_i}{s_i}) + C - \frac{1}{2}
\end{aligned}
$$

Removing constant terms we can get Equation M17.

1.5. Decomposition to sub-problems

The decomposition of the whole energy function (Equation M17) to sub-problems (Equation M20 is straightforward:

$$
\begin{aligned}
E_i &= \sum_{i=1}^{N_u} y_i^T \cdot (M_{ii}^{inter} + \lambda_1 M_i^{intra} + \lambda_2 \frac{e_i \cdot e_i^T}{s_i^2}) \cdot y_i + \sum_{i=1}^{N_u} \sum_{j=i+1}^{N_u} y_i^T \cdot M_{ij}^{inter} \cdot y_j + \sum_{i=1}^{N_u} y_i^T \cdot (V_i - \lambda_2 \frac{e_i}{s_i}) \\
&= \sum_{i=1}^{N_u} y_i^T \cdot M_i' \cdot y_i + y_i^T \cdot \sum_{j=1, j\neq i}^{N_u} M_{ij}^{inter} \cdot y_j + \sum_{j=1, j\neq i}^{N_u} y_j^T \cdot \sum_{k=j+1, k\neq i}^{N_u} M_{jk}^{inter} \cdot y_k + \sum_{i=1}^{N_u} y_i^T \cdot (V_i - \lambda_2 \frac{e_i}{s_i}) \\
&= y_i^T \cdot M_i' \cdot y_i + y_i^T \cdot \sum_{j=1, j\neq i}^{N_u} M_{ij}^{inter} \cdot y_j + y_i^T \cdot (V_i - \lambda_2 \frac{e_i}{s_i}) \\
&+ \sum_{j=1, j\neq i}^{N_u} [y_j^T \cdot M_j' \cdot y_j + y_j^T \cdot \sum_{k=j+1, k\neq i}^{N_u} M_{jk}^{inter} \cdot y_k + y_j^T \cdot (V_j - \lambda_2 \frac{e_j}{s_j})] \\
&= y_i^T \cdot M_i' \cdot y_i + y_i^T \cdot (\sum_{j=1, j\neq i}^{N_u} M_{ij}^{inter} \cdot y_j + V_i - \lambda_2 \frac{e_i}{s_i}) + C_i' \\
&= y_i^T \cdot M_i' \cdot y_i + y_i^T \cdot V_i' + C_i' \qquad (10)
\end{aligned}
$$

where M_i' and V_i' are defined in Equation M21 and M22 respectively and C_i' is the constant term for sub-problem E_i.

2. Experiment

We also show in Figure 1 the plots of the cosegmentation accuracy affected by the choice of parameters (λ_1 and λ_2), using "Baseball" class in iCoseg dataset as example. We can see that the accuracy would be stable when λ_1 is within a range of its best choice. For λ_2, increasing its value will increase the accuracy until reaching the best result, and further increasing it after the best choice would be slightly lower than the best accuracy but still stable. This results validate that the proposed method is not sensitive to the choice of parameters within a range.

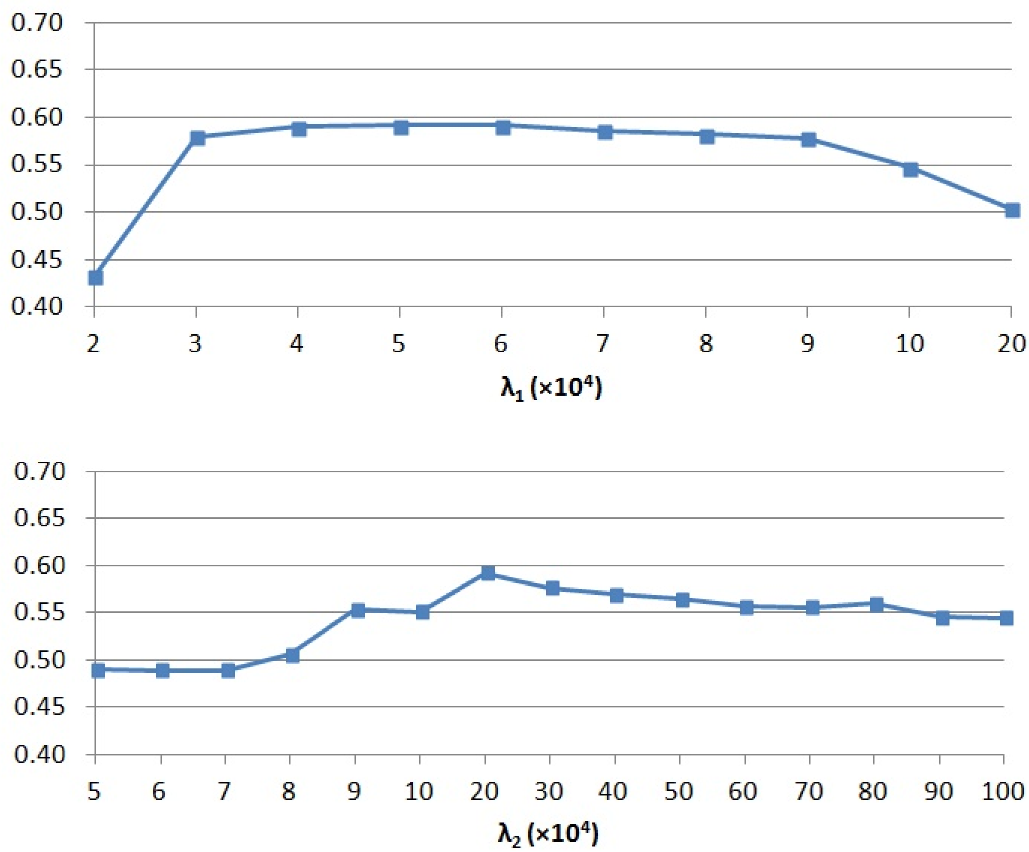

Figure 1. Accuracy under different parameter values for "Baseball" class in iCoseg dataset. Top: λ_1; Bottom: λ_2.