A Variational Co-training Framework for Remote Sensing Image Segmentation

Keming Chen[1], Zhenglong Li[1], Jian Cheng[1], Zhixin Zhou[12], Hanqing Lu[1]

[1]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China, 100190;

[2]Beijing Institute of Remote Sensing, Beijing, China, 100854

E-mail: {kmchen, zlli, jcheng, luhq@nlpr.ia.ac.cn}, zhixin.zhou@ia.ac.cn

ABSTRACT

Image segmentation is a fundamental process in remote sensing image interpretation. It is the basis of the image understanding, such as the region-based change detection for maps updating, the target recognition, and so on [3, 4]. This problem can be seen as a pattern classification application by employing a statistical framework, in which Bayesian inference has gained wide popularity in most recent research on computer vision. Bayesian inference is a theoretically well-founded and conceptually simple approach to data analysis. However, for practical interest, it is infeasible for exactly computations as the integrals appearing in Bayesian computations are seldom tractable. Variational Bayes (VB) is one of approximate inference techniques, the central idea of which is to fit a simple, tractable distribution to the posterior by variational methods [2].

Co-training, originally introduced in [1], works in a bootstrap mode. In co-strategy, each feature $X = \{x_i\}_{i=1}^{M*N}$ is divided into several disjoint subsets $X = (X^1, X^2, ..., X^j)$, and each subset $X^j = \{x_i^j\}_{i=1}^{M*N}$ is enough to learn the targets $f^j(x_i^j) \rightarrow y_i$. Such a subset is called a view. The essence of co-training is that: when initial estimates get stuck to local extrema in one view, other view can assist to pull it out.

Integrated the co-training strategy into VB approach [5], we propose a variational co-training framework (co-VB for short) for remote sensing image segmentation. Image data are characterized by different views and each view is modeled by a Gaussian mixture model (GMM). A factorized directed graph representation of the Bayesian mixture of Gaussian model under our co-VB framework is illustrated in *Fig. 1*.

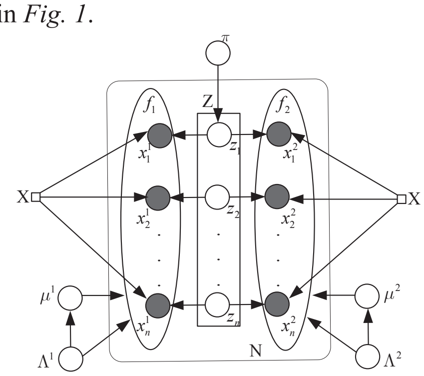

Figure 1. Two-view model for graphical representation of the Bayesian mixture of Gaussian model.

Given the observed data $X = \{x_i\}_{i=1}^{M*N}$ and the latent variable $Z = \{z_i\}_{i=1}^{M*N}$, in a Bayesian setting, a variational solution is to approximate $P(Z)$ with a simpler distribution $Q(Z)$ if it can provide a lower bound to the log-likelihood, $\mathcal{L}$(X) = *ln* P(X). We decompose the $\mathcal{L}$(X) = $\mathcal{L}$(q) + KL(q||p), where KL(q||p) is the *Kullback-Leibler* divergence. Inspired by the idea of co-training, we can formulate our segmentation model as following in a factorized distribution:

$$\begin{aligned} p(Z \mid X) &= p(X, Z, \Theta) = P(X \mid Z, \mu, \Lambda) p(\mu \mid \Lambda) p(\Lambda) p(Z \mid \pi) p(\pi) \\ &= p(X^1 \mid Z, \mu^1, \Lambda^1) p(\mu^1 \mid \Lambda^1) p(\Lambda^1) p(X^2 \mid Z, \mu^2, \Lambda^2) p(\mu^2 \mid \Lambda^2) p(\Lambda^2) p(Z \mid \pi) p(\pi) \end{aligned} \quad (1)$$

Where $p(\mu^j | \Lambda^j)$ and $p(\Lambda^j)$ are the conjugate prior. According to the factors $q(Z)=\prod_{i=1}^{K} q_i(Z_i)$ and $q(\Theta^j)$, for each pixel *i*, the distribution of the indicator z_{ik} for each component k is obtained:

$$\ln r_{ik} \propto (\psi(\alpha_k^1)-\psi(\hat{\alpha}^1))+(\frac{1}{2}\sum_{i=1}^{D}\psi(\frac{\nu_k^1+1-i}{2})+D\ln 2+\ln|W_k^1|)-\frac{1}{2}[D(\beta_k^1)^{-1}+\nu_k^1(x_n^1-m_k^1)^T W_k^1(x_n^1-m_k^1)] \tag{2}$$
$$+(\psi(\alpha_k^2)-\psi(\hat{\alpha}^2))+(\frac{1}{2}\sum_{i=1}^{D}\psi(\frac{\nu_k^2+1-i}{2})+D\ln 2+\ln|W_k^2|)-\frac{1}{2}[D(\beta_k^2)^{-1}+\nu_k^2(x_n^2-m_k^2)^T W_k^2(x_n^2-m_k^2)]$$

In order to thoroughly assess the co-VB performance, we designed two experiments on the basis of several different sets of high resolution optical remote sensing images, including Ikonos images, Quickbird images and Spot-5 images. Some segmentation results are shown in *Fig. 2*. In general, the experimental results confirm the effectiveness of the proposed algorithm. Co-VB method shows several advantages over classical single view based methods.

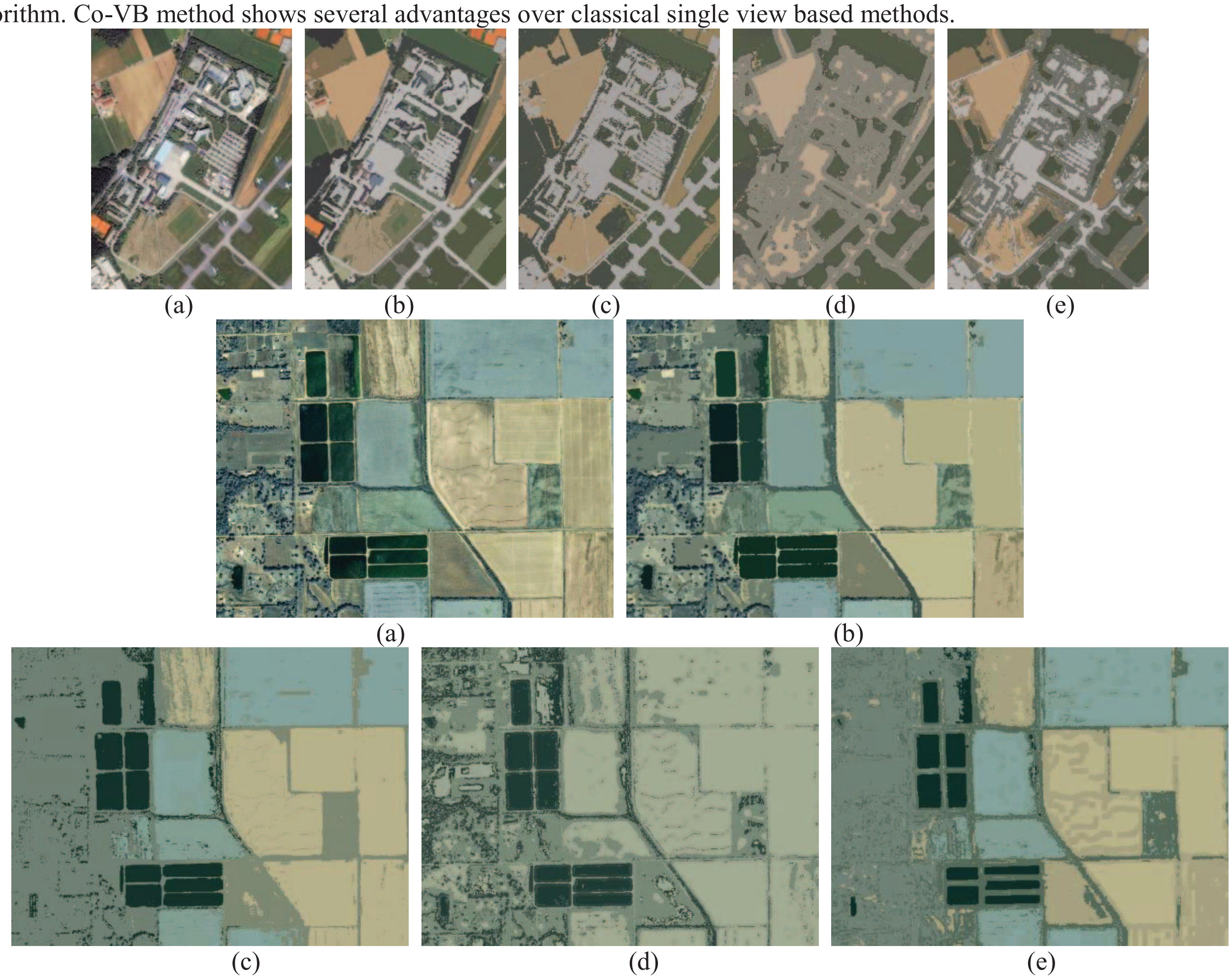

Figure 2. Visual comparison of some segmentation results. (a) original image; segmentation results obtained by (b) co-VB method; and single-view method with (c) color feature; (d) texture feature; (e) both color and texture feature.

REFERENCES

[1] A. Blum and Tom Mitchell, "Combining Labeled and Unlabeled Data with Co-Training," Proc. Conference on Computational Learning Theory, pp.92-100, July, 1998.

[2] C. M. Bishop, Pattern Recognition and Machine Learning. Springer, Heidelberg, 2006.

[3] M. Pesaresi and J. A. Benediktsson, “A new approach for the morphological segmentation of high-resolution satellite imagery”. IEEE Trans. Geoscience and Remote Sensing, Vol. 39(2), pp.309 – 320, Feb. 2001.

[4]Yashon O. Ouma, S.S. Josaphat and Ryutaro Tateishi. “Multiscale remote sensing data segmentation and post-segmentation change detection based on logical modeling Theoretical exposition and experimental results for forestland cover change analysis. Journal of Computers & Geosciences, 34(7), pp. 715-737, 2008.

[5] Zhenglong Li, Qingshan Liu, Jian Cheng and Hanqing Lu, “A Variational Multi-view Learning Framework and Its Application with Image Segmentation”. Technical Report, National Laboratory of Pattern Recognition, Chinese Academy of Sciences, 2008.