Image Segmentation using Dual Distribution Matching

Tatsunori Taniai[1]
taniai@nae-lab.org
Viet-Quoc Pham[2]
quocviet.pham@toshiba.co.jp
Keita Takahashi[3]
keita.takahashi@ieee.org
Takeshi Naemura[1]
naemura@nae-lab.org

[1] Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan
[2] Toshiba Corporate Research and Development Center, Kanagawa, Japan
[3] Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo, Japan

This paper addresses the problem of foreground-background image segmentation where only the approximate color distributions of the foreground and background regions are given as the input. Our aim is to derive a fundamental algorithm with this primitive setup that can find foreground and background regions that are consistent with the given input distributions. The essential question here is how to measure consistencies between the given distributions and the segmentation.

Local measures are widely adopted [2] by virtue of their simplicity. Each pixel is *individually* evaluated to determine how likely it is to belong to the foreground or background based on its color. However, local-measure-based methods are subject to the shrinking bias, which often results in shortcutting across thin structures.

Recent studies (*e.g.* BMGC [1]) have shown that methods based on *global measures* outperform conventional local-measure-based methods. The global consistency is measured by the similarity between a given distribution and the resulting distribution from the extracted region. We introduce a new distribution matching method named dual distribution matching (DDM) in order to increase the robustness of global measures. In this method, *the consistencies between two input distributions (the foreground and background distributions) and the resulting segmentation are enforced simultaneously*. Our method makes it possible to achieve robust and accurate segmentations even with not-so-accurate input distributions.

Dual Distribution Matching Binary segmentation is formulated as a problem that involves finding a label $\boldsymbol{L}$ for the set of pixels P, as $\boldsymbol{L} = \{L_p | L_p \in \{F, B\}, \forall p \in P\}$, where p denotes a pixel, and F/B denotes the foreground/background label. The foreground/background region is the set of all pixels with F/B and is denoted as $\mathbf{R}_l^{\boldsymbol{L}} = \{p \in P | L_p = l\}$ $(l = F, B)$. The probability distribution of colors (or intensities) within region $\mathbf{R}_l^{\boldsymbol{L}}$ is written as $\mathcal{P}_l^{\boldsymbol{L}}$ $(l = F, B)$.

Let us assume that only the approximate distributions for both the foreground and background are given as $\mathcal{H}_F \simeq \mathcal{P}_F^{\boldsymbol{L}^*}$ and $\mathcal{H}_B \simeq \mathcal{P}_B^{\boldsymbol{L}^*}$, where $\boldsymbol{L}^*$ is the ground truth of $\boldsymbol{L}$. Here, $\boldsymbol{L}^*$ is inferred as the label that minimizes the following energy function $\mathcal{E}(\boldsymbol{L})$:

$$\mathcal{E}(\boldsymbol{L}) = \underbrace{\lambda_F \mathcal{M}_F(\boldsymbol{L})}_{\text{Foreground Matching}} + \underbrace{\lambda_B \mathcal{M}_B(\boldsymbol{L})}_{\text{Background Matching}} + \underbrace{\lambda_S \mathcal{S}(\boldsymbol{L})}_{\text{Smoothness}}, \quad (1)$$

where $\mathcal{M}_l(\boldsymbol{L})$ is the negative of the distribution similarity measure $\mathcal{B}(,)$:

$$\mathcal{M}_l(\boldsymbol{L}) = -\mathcal{B}\left(\mathcal{P}_l^{\boldsymbol{L}}, \mathcal{H}_l\right) \quad (l = F, B). \quad (2)$$

The $\mathcal{S}(\boldsymbol{L})$ is a smoothness function composed of pairwise discontinuity penalties. This is called *dual distribution matching* or DDM, because both the foreground and background distributions are matched simultaneously. The term $\mathcal{B}(,)$ is the Bhattacharyya coefficient that measures the amount of overlap between two distributions f and g, which takes 1 as the maximum when $f = g$:

$$\mathcal{B}(f, g) = \sum_{z \in Z} \sqrt{f(z) g(z)} \le 1 \quad (3)$$

With the definitions above, $\mathcal{E}(\boldsymbol{L})$ with $\lambda_B = 0$ or $\lambda_F = 0$, which we define as $\mathcal{E}_F(\boldsymbol{L})$ or $\mathcal{E}_B(\boldsymbol{L})$ respectively, is equivalent to the single distribution matching of the BMGC method [1]. We refer to the BMGC method with $\mathcal{E}_F(\boldsymbol{L})$ or $\mathcal{E}_B(\boldsymbol{L})$ as F-BMGC or B-BMGC. As illustrated in Fig.1, those methods cannot capture the true solution $\boldsymbol{L}^*$ if the input distribution $\mathcal{H}_F$ or $\mathcal{H}_B$ is inaccurate. In contrast, *our method is more likely to capture the true solution by using both constraints simultaneously*. We show in this paper that $\mathcal{M}_F(\boldsymbol{L})$ *and* $\mathcal{M}_B(\boldsymbol{L})$ *should be weighted in proportion to the size of the foreground and background areas* so that the minimum solution of the energy function $\mathcal{E}(\boldsymbol{L})$ captures the true solution $\boldsymbol{L}^*$.

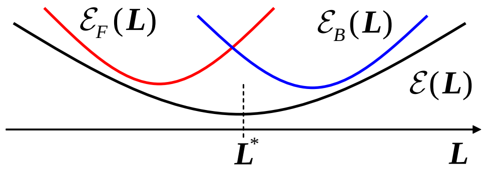

Figure 1: Dual matching energy function.

Experimental Results Figure 2 shows segmentation results of our method DDM, DDM with fixed weighting parameters, single distribution matching methods (F-BMGC and B-BMGC) [1], and local-measure-based method (interactive graph cuts) [2], where approximate input distributions are produced from foreground and background regions of lasso-trimap. Our method achieved the best accuracy in this experiment.

Target image & Ground truth	DDM (proposed)	DDM ($\lambda_F = \lambda_B = 0.5$)	F-BMGC [1] (foreground matching)	B-BMGC [1] (background matching)	Interactive graph cuts [2] (local)

Figure 2: Segmentation results with approximate input distributions.

Also, we compared local and global consistency measures while varying the accuracy of the input distributions. Figure 3 shows that *the proposed method outperforms the others at high and medium accuracies, whereas interactive graph cuts performed the best at very low accuracies*.

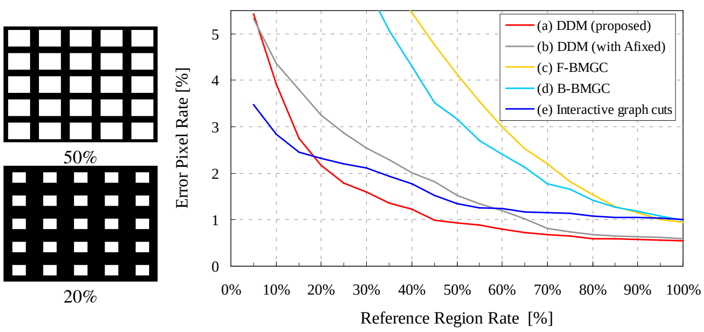

Figure 3: Comparison between local and global models according to input distribution accuracy. The input distributions were purposely made inaccurate by limiting the reference region using masks (left).

[1] Ismail Ben Ayed, Hua-Mei Chen, Kumaradevan Punithakumar, Ian G. Ross, and Shuo Li. Graph cut segmentation with a global constraint: Recovering region distribution via a bound of the bhattacharyya measure. In *Proc. CVPR*, pages 3288–3295, 2010.

[2] Yuri Boykov and Marie-Pierre Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. In *Proc. ICCV*, pages 105–112, 2001.