

CV Project Proposal: Image Parsing

103062588 呂立嫻

1. Introduction

Image parsing 的目的是對一張影像中的每一個 pixel 給予它們的 label，而這個 label 可以正確解釋這些 pixel 代表的意義。例如有一張街景的照片，透過 Image parsing 可以將畫面中的建築物、車子、街道、行人、天空等，皆給予正確的 label 並視覺化結果 (Figure 1)。



Figure 1. Example of image parsing. (Left: Input image\ Right: Final parsing).

這項技術可以延伸到如 google glasses 的裝置上面，使用者偷過眼鏡看到景象，glasses 可以分析景物的內容並標示出來，即時顯示在螢幕上面。

2. Technical Part

計畫結合 region-level features 以及 per-exemplar sliding window detectors 來實作 image parsing。過去的作法大多使用 bounding box detectors，和 bounding box detectors 相比，per-exemplar detector 在處理 training data 較小以及高 intra-class variation 的 case 時能達到不錯的效果，且在做 pixel-level segmentation 時可以將 object mask 套用到 test image 上。

作法如 Figure 2 所示，讀入一張影像，分成兩個部分，第一部分為 region-based parsing，第二部分為 detector-based parsing，經過兩個部分的 parsing 各可以得到一個初步的結果，但效果可能不是那麼好，於是再將兩個結果結合得到最後較好的結果。

Training part 計畫對每一個 class 都 train 一個 one-vs-all 的 SVM 來對每一個 pixel 做分類。

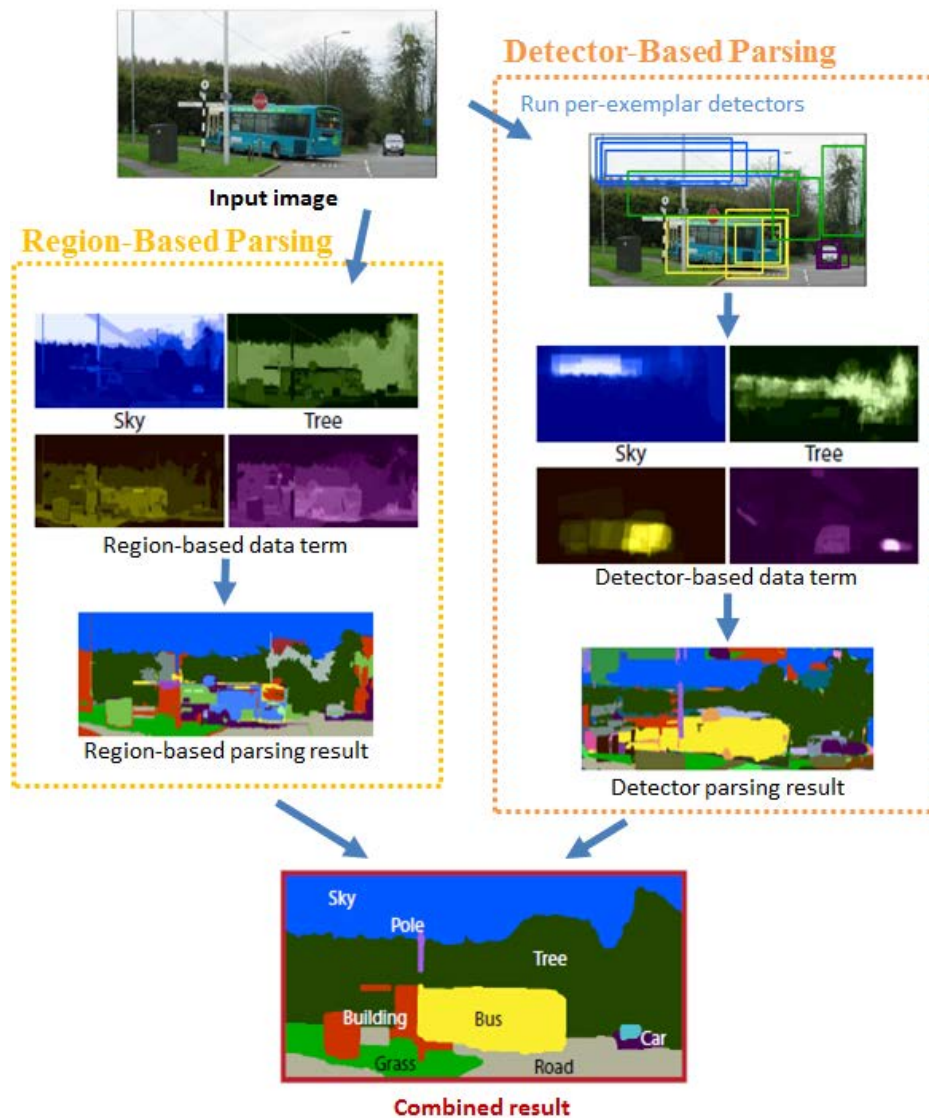


Figure 2. Flow chart.

3. Milestones

- 2014.11.30 Region-based parsing
- 2014.12.11 Per-exemplar detector + midterm project report
- 2014.12.30 Detector-based parsing
- 2015.01.11 Combine two parsing parts

4. References

- [1] J. Tighe and S. Lazebnik. SuperParsing: Scalable nonparametric image parsing with superpixels. *IJCV*, 101(2):329–349, Jan 2013.
- [2] T. Malisiewicz, A. Gupta, and A. A. Efros. Ensemble of exemplar-SVMs for object detection and beyond. In *ICCV*, 2011.
- [3] J. Tighe and S. Lazebnik. Finding things: Image parsing with regions and per-exemplar detectors. In *CVPR*, Jun 2013.