COMP2550/COMP3130 ANU Main Project Proposal

Christopher Claoué-Long (*u5183532@anu.edu.au*)
Jimmy Lin (*u5223173@anu.edu.au*)
Last typeset: 21:06, Wednesday 17th April, 2013

1 Introduction

[problem motivation, background, application...]

There are a few existing works to achieve the functionality of detecting the salient objects. The mainstream appraoch in the last decade should be the (Itti,1998)'s algorithm (see ref.3). It outputs a feature map and then converts it to a rectangle by winner-take-all algorithm. But the precision of the detection is quite unsatisfying, even though it does hae a good recall. The second approach is the fuzzy-growing approach (see ref.4). But [badness of this approach...].

(a) FG(Ma,2003) (b) SM(Itti,1998) (c) CRFM(Liu,2007) (d) Ground truth

The research we referred to in this experimental project (see ref.2), which is based on conditional random field (CRF) model, has a large portion of perfect detection compared with ground truth data. At first stage, it extracts features in the local, regional and global level, corresponding to multiscale contrast, center-surround histogram and spatial color distribution respectively. And then after normalization and linear/non-linear combination, a master map or a salient map is computed to represent the saliency of each image pixel. Last, a few key locations on the saliency map are identified by winner-take-all or inhibition-of-return, or other non-linear operations.

Our implementation involves in utilizing the open-source library OpenCV and Darwin to achieve the framework of the CRF-based saliency detection. [why we choose this topic, (1) experience of implementing graphical model (2) it may help to scene understanding...]

2 Formulation

2.1 CRF Model

Model conditional distribution of objective variable

$$P(A|I) = \frac{1}{Z} exp(-E(A|I))$$

Energy function is formualted to be a set of static salient features and one pairwise feature as follows,

$$E(A|I) = \sum_{x} \sum_{k=1}^{K} \lambda_k F_k(a_x, I) + \sum_{x,x'} S(a_x, a_{x'}, I)$$

λ_k: weight of kth feature, x, x': two adjacent pixels.

2.2 Feature Extraction

Multiscale Contrast. This static feature captures the high contrast in the boundary of objects .

Center-Surround Histogram. This static feature captures

Spatial Color distribution. This static feature penalizes the pixels with widely distributed color.

Pairwise Feature. This feature exploits the spatial relationship between two adjacent pixels

and can be viewed as one capturing the spatial continuity of saliency, that is, adjacent pixels that are prone to be assigned with different labels.

2.3 Possible Improvements

[more features that better capture the saliency...]

3 Project Timeline

Apr. 1st ⇒ Apr. 18th
Project Proposal. Read papers to find project of our interest, determine the topic of our second project and collect relevant data sets for training.

Apr. 19th ⇒ Apr. 28th
Framework Construction. Familiarize ourself with the related packages in Darwin and OpenCV, set up the interface to accept training data and the framework of CRF model for learning and inference.

Apr. 29th ⇒ May. 12th
Intensive Coding. Write codes to implement extraction of various features, and design the algorithm to output a rectangle labeling saliency.

May. 12th ⇒ May. 19th
Testing, Enhancement. Test the framework we would have constructed. Search and implement possible improvements.

May. 20th ⇒ May. 30th
Project Summarization. Write report to summarize our project and make presentation.

References

[1] Gould, Stephen. "DARWIN: A Framework for Machine Learning and Computer Vision Research and Development." *Journal of Machine Learning Research 13 (2012): 3533-3537*.

[2] Liu, Tie, et al. "Learning to detect a salient object." *Pattern Analysis and Machine Intelligence, IEEE Transactions on 33.2 (2011): 353-367*.

[3] Itti, Laurent, Christof Koch, and Ernst Niebur. "A model of saliency-based visual attention for rapid scene analysis." *Pattern Analysis and Machine Intelligence, IEEE Transactions on 20.11 (1998): 1254-1259*.

[4] Ma, Yu-Fei, and Hong-Jiang Zhang. "Contrast-based image attention analysis by using fuzzy growing." *Proceedings of the eleventh ACM international conference on Multimedia. ACM, 2003*.