Paper Commentary Exercise

Janet Huang

October 31, 2014

1 Scalable Multi-label Annotation – Rating: 4/5

Deng *et al.* [1] address the scalablity issue of multi-label annotation. They proposes an algorithm that exploits correlation, hierarchy, and sparsity of the label distribution to yield significant savings for image labeling tasks.

The paper clearly identifies three key obersvations for labels in real world scenarios. The labels are correlated, sparse and naturally form a hierarchy.

The authors propose a theoretical analysis and a practical algorithm to satify the assumptions. They clearly define the utlity and cost function and use a pseudo code to describe the algorithm.

The paper applys the algorithm on a real image dataset in real world scenarios. The results show that their algorithm accuires up to 6 savings compared to the naïve approach.

I wish they can provide a complete example to describe the process of anaylysis and demostrate the power of algorithm for reducing the costs. To achieve the goal of scalablity, we should select the high-utility queries with low cost. I wish the authors can provide their ideas on the trade-off between utility and cost.

References

[1] J. Deng, O. Russakovsky, J. Krause, M. S. Bernstein, A. Berg, and L. Fei-Fei. Scalable multi-label annotation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 3099–3102, New York, NY, USA, 2014. ACM.