

Local Feature Based Salient Region Detection

Chao Lv
Parallel Processing Institute
Fudan University
lch@fudan.edu.cn

ABSTRACT

Local feature descriptors have become the most important part in image / video retrieval systems. But considering the great amount of local features, which could be thousands of local features in one HD photo, it's hard to compute them efficiently in a realistic system. In our research, we try to overcome this obstacle with a straightforward local feature reduction processing by using salient region detection. Furthermore, we also present a efficient method to detect multi salient regions. The whole algorithm is only based on the local features, which means almost no additional overhead involved in our algorithm. In our tests, we compare the LFSR (Local Feature based Salient Region) algorithm with a state-of-the-art algorithm. And the LFSR shows a thousands of times speedup in runtime with acceptable precision and recall loss. When integrated with the SURF descriptor, LFSR can provide a overall 1.6X speedup for the whole processing [?].

1. INTRODUCTION

Our society has entered a data-centric era and a huge amount of data are transferred and processed on the Internet. Among them, multimedia data, such as image and video, has become one of the major data types being processed. As analyzed by CISCO Inc., video data occupies 50% of network traffic in 2011 and will increase to 90% in 2013 [1]. According to a report [6], as one of the most popular video sharing sites, more than 20-hour new videos are uploaded to *YouTube* every minute. Moreover, as two most popular photo sharing sites, *Facebook* and *Flickr* host billions of user-uploaded images respectively.

With the rapid increase of multimedia data, one of the most significant challenges is to understand and interpret such a huge amount of multimedia data. Currently, more and more retrieval applications are emerging to process these multimedia data, such as video recommendation [10], travel guidance systems [5] and content-based TV copy identification [7]. In these systems, a fundamental step is to extract

feature information from images.

Image features can be divided into two domains – local features and global features. Local features are more precise and robust when comparing to global features. Thus they are widely used in most of image retrieval systems. But the computation of local features suffers great overhead due to thousand of features existed in one image, which leads to a great challenge for real-time processing. For example, SIFT [8] and SURF [3], two popular image retrieval algorithms [9] [2], can only achieve a speed of about three images per second, when the computation are processed on a general-purpose processor.

In general, there exist three major computation phases when processing local image features in a typical image retrieval system. First, the system detect all feature points from images. Then, with some specific formats and algorithms, each feature point is described as a high dimension vector. At last, all extracted feature are compared to each other according to the distance of their vectors.

According to a previous research [4], the computation of feature describing are obviously greater than the feature detecting in the SURF algorithm, which is caused by the great amount of local features to be described in one image. Furthermore, in a realistic image retrieval system, the performance is dominated by the number of features in the database.

So, it's possible and necessary to improve the performance of the whole system by the reduction of local features extracted from each image. In this paper, we present a algorithm named LFSR (Local Feature based Salient Region) to eliminate unimportant local features efficiently with no obvious precision loss for a image retrieval system.

The main idea of LFSR algorithm is extracting salient regions of images and only describing the local features in those regions. Without involving any other salient region algorithm, it is only based on the feature points extracted from the first phase of a typical local feature algorithm. This approach has two remarkable advantages: no additional computation for salient region detection; totally integrated with local feature algorithm to compute the salient region efficiently. The details of LFSR can be found in Section 2.

We have evaluated LFSR against a stat-of-the-art salient

region algorithm. The evaluation results show that our approach has a much better performance with comparable precision and recall. Furthermore, withing a realistic image retrieval system, our evaluation shows that LFSR can help to improve both performance and accuracy.

2. CONCLUSION

Acknowledgement

3. REFERENCES

- [1] Cisco Visual Networking Index: Forecast and Methodology, 2010-2015. 2011.
- [2] J. Bauer, N. Sunderhauf, and P. Protzel. Comparing Several Implementations of Two Recently Published Feature Detectors. *International Conference on Intelligent and Autonomous Systems*, 2007.
- [3] H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded up robust features. *European Conference on Computer Vision*, pages 404–417, 2006.
- [4] Z. Fang, D. Yang, W. Zhang, H. Chen, and B. Zang. A Comprehensive Analysis and Parallelization of an Image Retrieval Algorithm. *ISPASS*, 2011.
- [5] Y. Gao, J. Tang, R. Hong, Q. Dai, T. S. Chua, and R. Jain. W2GO: a travel guidance system by automatic landmark ranking. *ACM Multimedia*, pages 123–132, 2010.
- [6] C. Jansohn, A. Ulges, and T. M. Breuel. Detecting pornographic video content by combining image features with motion information. *International Conference on Multimedia*, 2009.
- [7] A. Joly, C. Frelicot, and O. Buisson. Robust Content-Based Video Copy Identification in a Large Reference Database . *International Conference on Image and Video Retrieval*, pages 511–516, 2003.
- [8] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [9] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, 2005.
- [10] B. Yang, T. Mei, X.-S. Hua, L. Yang, S.-Q. Yang, and M. Li. Online video recommendation based on multimodal fusion and relevance feedback. *International Conference on Image and Video Retrieval*, pages 73–80, 2007.