Research Proposal

Contact person: Prof. Ming-Ming Cheng (Homepage: http://mmcheng.net/)
Address: College of Computer and Control Engineering, Nankai University, Tianjin, China. 300071

Short description of research projects:

I'm currently working on understanding the visual word and editing the visual media. These works are mainly in following aspects: (I) efficient algorithms for scene understanding; (II) realtime 3D reconstruction and understanding; (III) interactive image manipulation. These works tried to understand the visual scene in realtime while allowing users to be in the loop if required.

With collaborators from Oxford University, Microsoft research Cambridge, Tsinghua University, etc., we have started some of the cutting edge technologies which use parallel algorithms. For object detection proposal generation task, our IEEE CVPR 2014 oral paper [4] achieves best results in PASCAL VOC benchmark while being 1000 times faster than previously state of the art. In the area of 3D understanding, we have developed the world's first realtime 3D fusion, online learning, physical interaction, and realtime classification system (accepted by ACM TOG 2015). In the area of salient object detection, our algorithm is widely used in many area of research and industry.

In the past 5 years, I have published about 20 papers which highly influenced the research community as indicated by the 2000+ citations to my papers, and numerous international awards I have received. Our Sketch2Photo video in Vimeo has attracted 1,000,000+ clicks within 6 months and was selected as one of the "10 most innovative and promising worldwide initiatives of 2009" by the Netexplorateur jury, under the aegis of French government. Our C++ source code have received 8000+ downloads according to the records of email addresses.

Plan of using the GPU:

As described above, our research work have been focused on efficient algorithms which has been benefit from the parallel computing ability enabled by Nvidia GPUs [1, 3, 5, 6]. We plan to use the latest NVidia GPU to push our research a big step forward, while continuously making our algorithms open source. We believe the surprise efficiency of our algorithm and its open source CUDA/C++ code, will largely expand the applications and usages of NVidia GPUs.

List of recent publications:

[1] SemanticPaint: Interactive 3D Labeling and Learning at your Fingertips, ACM TOG 2015.
[2] Global Contrast based Salient Region Detection, IEEE TPAMI 2015. (CVPR 2011 version get 700+ citations)
[3] ImageSpirit: Verbal Guided Image Parsing, ACM TOG 2014.
[4] BING: Binarized Normed Gradients for Objectness Estimation at 300fps, IEEE CVPR (Oral) 2014.
[5] Dense Semantic Image Segmentation with Objects and Attributes, IEEE CVPR 2014.
[6] Efficient Salient Region Detection with Soft Image Abstraction, IEEE ICCV 2013.
[7] PoseShop: Human Image Database Construction and Personalized Content Synthesis, IEEE TVCG 2013.
[8] Interactive Images: Cuboid Proxies for Smart Image Manipulation, ACM TOG 2012.
[9] Connectedness of Random Walk Segmentation, IEEE TPAMI 2011.
[10] RepFinder: Finding Approximately Repeated Scene Elements for Image Editing, ACM TOG 2010.
[11] Sketch2Photo: Internet Image Montage, ACM TOG 2009.

Notes: **IEEE TPAMI** is the #1 IEEE publication. **ACM TOG** is the most-cited ACM Transaction. **SIGGRAPH**, **CVPR**, and **ICCV** have CiteSeer impact factor ranking top 0.7%, top 5%, and top 5% of all computer science journals and conferences respectively.