

# CHUAN WANG

Specialise in Computer Vision and Image/Video Processing

(+852) 5162 6756 (HK)  
(+86) 131 4751 1986 (China)  
✉ [wangchuan2400@gmail.com](mailto:wangchuan2400@gmail.com)  
📄 [wangchuan.github.io/cv](https://wangchuan.github.io/cv)

## — OBJECTIVE —

I am seeking a position in IT industry where I can contribute my proficient programming skills, energy, high motivation and dedication to the company. My research interests include but not limited to image/video processing, 3D vision, deep learning and graphics.

## — EDUCATION —

Ph.D, Computer Science (Vision and Graphics), 2015, *The University of Hong Kong (HKU)*.

B.Eng, Electronic Information Engineering, 2010, *University of Science and Technology of China (USTC)*.

## — WORK AND RESEARCH EXPERIENCES —

**Staff Researcher, Lenovo** (April. 2015 - Now)

### ○ **An Augmented Reality Application for Mobile Devices.**

- Developed an application providing augmented reality effect when camera previewing scenes containing QR code or a dish of food, to improve user experiences.
- The application detects and tracks the features as well as estimates the camera poses by the 3D-2D perspective n-points correspondence algorithm. With the camera pose estimated for each frame, we render an virtual object by OpenGL for the scene.
- Designed a strategy to switch between detecting and tracking the features, in order to obtain a real-time and accurate performance with smooth transition of camera pose.

### ○ **An Image Refocus Algorithm for Dual-camera Cellphone of Lenovo.**

- Developed an image refocus algorithm for Lenovo dual camera phone [VIBE S1](#), which is on sale in the market.
- The algorithm uses the depth information captured by the dual camera to blur the image with a spatially-variant kernel if users specify a focus point when viewing the image.
- To achieve real-time interaction, we utilized OpenCL to speed up the blurring process. In addition to OpenCL, we also utilized block-wise processing idea to balance between speed and memory cost.

### ○ **Speaker Identification via Deep Learning.**

- Co-developed a speaker identification system based on Recurrent Neural Network (RNN), specifically Long Short-Term Memory (LSTM) technique.
- The system utilizes both face image and audio information in a video sequence to classify the speaker. The novelty lies in such a multi-modal LSTM solution.
- Did experiments to demonstrate this solution can achieve higher precision than state-of-the-art Convolution Neural Network (CNN) based method, implemented based on Caffe. The paper was accepted in AAAI 2016.

**Ph.D, The University of Hong Kong** (Sep. 2010 - Jan. 2015)

### ○ **Video Vectorization.**

- Proposed a vector-based representation, i.e. tetrahedral mesh, for raster video to achieve low storage with resolution-independent high quality of reconstruction. Also developed a system to convert a raster video into its vectorized version.
- Developed a core algorithm for the system, i.e. a tetrahedral mesh simplification algorithm extended from the one for triangular mesh but handles more complicated topological cases. Also implemented subdivision and deformation algorithm for the tetrahedral mesh to preserve the features in the original video.
- Developed a rasterization method to convert the vectorized version back, i.e.
  - slice the tetrahedral mesh into triangular meshes;
  - render each triangular mesh as each frame by OpenGL.
- Designed a streaming strategy to process a video to achieve low memory cost but keeps the topology of the corresponding tetrahedral mesh.

### ○ **Video Object Co-Segmentation.**

- Developed a common-foreground co-segmentation system for a group of videos automatically.
- Proposed a novel customized subspace clustering algorithm to differentiate foreground and background within each video but correlate the common foreground in various videos. This algorithm takes advantage of appearance and motion features in the video group discriminatively, so as to outperform state-of-the-art methods.
- To build up the system, utilized existing mature techniques like K-means algorithm, Gaussian Mixture Model, Bag-of-Words Model and Markov Random Field etc. The paper was published in TMM 2014.

### ○ **Image Matting Application.**

- Developed an interactive application to cut foreground from an image based on Robust Image Matting algorithm.

### Teaching Assistant, The University of Hong Kong (Sep. 2010 - Dec. 2013)

- TA for courses Image Processing and Computer Vision, Multimedia Technologies, etc.
- Helped students master knowledge in Fourier analysis, machine learning algorithms and related image/video applications; Taught students image programming in MATLAB and C++.

### Research Intern and Developer, State Key Lab of CAD&CG, Zhejiang University (Mar. 2010 - Jun. 2010)

- Developed a segmentation and visualization application for medical images based on Graph Cuts algorithm and volume rendering technique.

## SKILLS

### Programming

- Proficient in Native C/C++, MATLAB.
- Proficient in OpenCV; Familiar with OpenGL; Experience of using OpenCL, CUDA, Boost, Intel MKL, CGAL etc.
- Proficient in developing with Visual Studio in Windows; Experience of developing in Linux.
- Familiar with using computer vision or machine learning related techniques such as CNN(Caffe), Optical/SIFT Flow, Image Deblurring, Calibration, Structure from Motion etc.
- Familiar with .NET Framework, C#, WPF, Managed C++.
- Familiar with GUI design in WPF or QT; Love designing applications with fancy GUI (e.g. [here](#) by C++/C#/WPF).

### Tools

- Familiar with  $\LaTeX$ , Markdown, Microsoft Office, Adobe Photoshop/Illustrator/Premiere, MeshLab.

### Languages

- Full professional proficiency in English.
- Native in Chinese Mandarin.

## SELECTED PUBLICATIONS

- **Chuan Wang**, Yanwen Guo, Jie Zhu, Linbo Wang, Wenping Wang. Video object co-segmentation via subspace clustering and quadratic pseudo-boolean optimization in an MRF framework. IEEE Transactions on Multimedia 2014, 16(4): 903-916.
- Jimmy S.J. Ren, Yongtao Hu, Yu-Wing Tai, **Chuan Wang**, Li Xu, Wenxiu Sun, Qiong Yan, "Look, Listen and Learn - A Multimodal LSTM for Speaker Identification", The 30th AAAI Conference on Artificial Intelligence (AAAI-16).

*Other TIP and TMM submissions are under review.*