

WEICAI YE

✉ yeweicai@zju.edu.cn · ☎ (+86) 130-3289-3650 · in weicaiye · 🌐 ywcmaike · 📄 ywcmaike.github.io

🎓 EDUCATION

Zhe Jiang University (ZJU), Hangzhou, China 2018 – Present

Master student in Computer Science (CS), expected March 2021

University of Electronic Science and Technology of China, Chengdu, China 2014 – 2018

B.S. in Computer Software Engineering (CSE)

👥 PROJECTS AND PUBLICATIONS

Learning Bipartite Graph Matching for Robust Camera Localization Jun. 2019 – Nov. 2019

Hailin Yu, Weicai Ye, Youji Feng, Hujun Bao, Guofeng Zhang

- Feature matching is an important step for many computer vision applications. The follow up is determined by the quality of correspondences established by feature matching, which heavily depends on the representation power of local features. Recently, due to the powerful representation of deep learning, traditional handcrafted local features have been replaced by CNN-based ones, which show better results in some complex scenes. However, it is still difficult to establish enough correct matches in scenes with local illumination changes or repeated patterns due to using local similarity. In this work, we propose a novel method to address the problem which first establishes as many potential correct matches as possible using local similarity, then utilizes global geometric information to reject ambiguous matches and generate globally optimal one-to-one matches. To achieve this, we propose a Bipartite Graph Network, called BG-Net then applied it to the camera localization problem. Extensive experiments demonstrate that our method can establish more correct matches, thereby improving the accuracy of the solved pose, and achieving state-of-the-art results on multiple visual localization datasets.
- Rank 3rd with pure BG-Net, Long-Term Visual Localization Challenge, CVPR2020 Workshop. The whole method won on multiple visual localization benchmark.
- The paper will be submitted to ISMAR2020.
- The code will be available at: <https://github.com/ywcmaike/BipartiteGraphNet>.

Saliency Guided Subdivision for Single-View Mesh Reconstruction Sep. 2019 – Jan. 2020

Hai Li, Weicai Ye*, Guofeng Zhang*

- In this paper, we present a novel deep architecture to recover a 3D shape in triangular mesh from a single image based on mesh deformation. Most existing deform-based methods produce uniform mesh predictions by repeatedly applying global subdivision but fail to require the highlighted details due to the memory limits. In order to address this problem, we propose a novel saliency guided subdivision method to achieve the trade-off between detail generation and memory consumption. Instead of using local geometric cues such as curvature, we introduce a point-based saliency voting operation to guide the adaptive mesh subdivision and deformation explicitly. Moreover, we introduce an orient chamfer loss to address the mesh self-intersection problem in subdivision. We further make our network configurable and explore the best structure combination. Extensive experiments show that our method can not only produce visually pleasing results with fine details but also achieve better performance compared to other state-of-the-art methods.
- The paper will be submitted to MM2020.
- The code will be available at: <https://github.com/ywcmaike/SaliencyGuidedSubdivision>.
(* indicates equal contribution)

SuperPlane: 3D Plane Detection and Descriptor from a Single Image Sep. 2019 – Present

Weicai Ye, Guofeng Zhang

- Human-made environments usually contain many planar structures, which are often weak or repeating textures. This makes it difficult for feature-based methods to extract robust features or tend to cause mismatches. To tackle this problem, this paper proposes a new task—detecting 3D planes and descriptor gener-

ation from a single image. We propose a novel baseline named SuperPlane that contains a mask attention module to capture the plane feature. We also create plane descriptor benchmarks to verify the effectiveness of our baseline. In addition, we consider planes as feature points and apply them to pose estimation tasks. Extensive experiments demonstrate that the proposed method achieves comparable results in a general environment, and outperform in weak or repeat-texture scenes than the feature point-based methods. With the descriptor, the plane segmentation and depth estimation have been significantly improved. We highlight that the mid-level feature can be applied to the Visual Localization tasks.

- Extra experiments have been conducted and the paper is writing now.
- The code will be available at: <https://github.com/ywcmake/SuperPlane>.

Cross-Modal Shop Clothes Retrieval for Live Delivery

Mar. 2020 – Present

Weicai Ye, Delv Lin, Qi Chen, GuoFeng Zhang

- Live streaming is an important way for Taobao to connect goods with consumers. Buyers buy their favorite products while watching the live broadcast. In a single Taobao live broadcast, the anchor will often display, try, and introduce hundreds of products. If a buyer wants to purchase the product being explained, they need to list the products associated with the live broadcast to manually select, which greatly affects the user's purchasing efficiency and user experience. To tackle this problem, we propose a novel one-stage clothes retrieval framework which directly to acquires the clothes' location and ID embedding. A motion capture module is proposed to effectively identify the anchor's intentions from the live broadcast and retrieve products related to the live broadcast merchandise. In order to exploit text information, we propose a cross-domain fusion method. Then a rerank post-process is followed to further improve retrieval accuracy. Extensive experiments show that our algorithm achieves the competitive results compared to the detect-match approaches, and will be applied to the industrial environment of Taobao live broadcast with goods to achieve a balance between accuracy and performance.
- The one-stage baseline is test on Tianchi and now ranked 6th. The improve version such as motion attention, cross-domain fusion and rerank postprocess have been conducted, but not tested.
- We also conduct the detect-match approaches as comparison to verify the effectiveness of our novel one-stage methods.
- The technique report will be written then submitted to CVPR2021.
- The code will be available at : <https://github.com/ywcmake/ClothesRetrieval>.

Cloudwalk Headcount Challenge

Nov. 2018 – Jan. 2019

Weicai Ye, Chunfang Deng, Rang Meng

- Modify the open source code framework of the face detection paper "pyramid box", add a priori such as shoulders and upper body to the model to improve the accuracy of the model in crowded people, and use multi-scale information fusion to improve the accuracy of small objects.
- Warm up the learning rate during training, and learn rate decay at specific iterations. Utilize a variety of tricks in testing, such as flip testing, multi-scale testing, bounding boxes voting, multi-model ensemble, etc to improve the accuracy of head detection.
- Champion of 2018 Cloudwalk Headcount Challenge with 31,500¥ Bonus.
- The code is available at : <https://github.com/ywcmake/2018-ZJUI-PyramidBoxDetector>.

EXPERIENCE

SenseTime Group Ltd. Hangzhou, China

Jan. 2018 – May. 2018

3D Vision Researcher Internship

3D Reconstruction of Indoor Scene of RGB-D Images

- Integrate traditional RGBD SLAM and semantic segmentation, plane detection to form SemanticSLAM.
- Different modules cooperate with each other to effectively improve the quality of localization and mapping.
- The system can perform real-time reconstruction of room-scale indoor scenes.

Baidu Inc. Beijing, China

Feb, 2017 – Aug, 2017

Software Engineer Internship

Video Search: Given the text and Retrive the related video.

- Cooperate with colleagues to develop millisecond-level response video search services that can support hundreds of millions of highly concurrent retrieval needs.
- Cooperate with colleagues to develop rearrangement strategies such as video resolution, cross-modal fusion, etc to improve the quality of video retrieval results.
- The system can retrieve the related video with the given text in real time and the service is online.

♥ HONORS AND AWARDS

Zhijun He Outstanding Scholarship.	2019
Chiang Chen Industrial Charity Foundation Grant.	2019
Excellence Price of the 1st IKCEST "The Belt and Road" International Big Data Competition.	2019
Champion of 2018 Cloudwalk Headcount Challenge with 31,500¥ Bonus.	2019
Graduate Student Scholarship.	2018,2019,2020
National Encouragement Scholarship, Ranked 3rd of 111 Students.	2017
Meritorious Winner In Mathematical Contest In Modeling.	2017
1 st Prize in Sichuan Province Contest District in China Undergraduate Mathematical Contest in Modeling.	2016

📖 TEACHING ASSISTANT

- Introduction to Python Programming, Teaching Assistant (TA) - Spring 2020
- Introduction to Java Programming, Teaching Assistant (TA) - Spring 2019

⚙️ SKILLS

- Programming Languages: Python/C,C++/Java/Shell/CUDA
- Tools: Pytorch, Git, Latex
- Development: Web Development
- Languages: English - Fluent, Chinese - Native speaker