Research Statement

I am a fourth year Ph.D. student in the Bradley Department of Electrical and Computer Engineering at Virginia Tech. I work in the Vision and Learning Laboratory headed by Prof. Jia-Bin Huang. My research interests are in the area of computer vision and computer graphics. I am particularly interested in realistic controllable image synthesis.

Education

Virginia Polytechnic Institute and State University (Virginia Tech)
PH.D. candidate in COMPUTER ENGINEERING

Virginia Polytechnic Institute and State University (Virginia Tech)
M.S. in COMPUTER ENGINEERING
Thesis: "Im2vid: Future Video Prediction for Static Image Action Recognition".

Kuwait University
B.S. in COMPUTER ENGINEERING
Distinction with Class Honors

Publications

CONFERENCE PROCEEDINGS
Guided Image-to-Image Translation with Bi-Directional Feature Transformation
Badour AlBahar, Jia-Bin Huang
Proceedings of the IEEE International Conference on Computer Vision, 2019

JOURNAL ARTICLES
Pose with Style: Detail-Preserving Pose-Guided Image Synthesis with Conditional StyleGAN
Badour AlBahar, Jingwan Lu, Jimei Yang, Zhixin Shu, Eli Shechtman, Jia-Bin Huang

Work Experience

Adobe Research
RESEARCH INTERN
San Jose, California
Summer 2021

Research Experience

June 2020 - May 2021  Pose with Style
In this work, we re-render a person from a single image under arbitrary poses. Existing methods often have difficulties in hallucinating occluded contents photo-realistically while preserving the identity and fine details in the source image. We first learn to inpaint the correspondence field between the body surface texture and the source image with a human body symmetry prior. The inpainted correspondence field allows us to warp local features extracted from the source to the target view even under large pose changes. To generate photo-realistic results, we extend the StyleGAN generator so that it takes pose as input (for controlling poses) and introduce a spatially varying modulation for the latent space using the warped local features (for controlling appearances).
Why do padding? How does it effect image processing?

During this time, we have focused on the effect of padding in deep learning. Why use zero/reflection padding? and why does partial convolution perform better? Can we better handle the boundary? and how would gated convolution handle boundaries if we were to consider them as "holes" or missing regions. We have come up with several approaches that mainly focus on treating the boundary as missing region and tried to fix up the computation involving those regions. More details of this work can be found as part of my colleague’s thesis here.

Guided Image-to-Image Translation

In this work, we address the problem of guided image-to-image translation, where we aim to translate an input image into another while respecting the constraints provided by an external, user-provided guidance information. There are various conditioning schemes that can be used to incorporate this given guidance signal, including input/feature concatenation and conditional affine transformation of feature activations. In this work, we present bi-directional feature transformation (bFT), a new conditional scheme for guided image-to-image translation problems. Our core technical contributions lie in the use of spatially varying feature transformation and the design of bi-directional conditioning scheme that allows mutual modulation of the guidance and input network branches. We validate the applicability of our method on various tasks. While being application agnostic, our approach achieves competitive performance with state-of-the-art methods. The generality of our method opens promising direction of incorporating a wide variety of constraints for image-to-image translation problems.

Future Video Prediction and Static Image Action Recognition

In this work, we address the problem of static image action recognition, which aims at identifying the action performed in a given image. Most existing static image action recognition approaches use high-level cues present in the image such as objects, object human interaction, or human pose to better capture the action performed. Unlike images, videos have temporal information that greatly improves action recognition by resolving potential ambiguity. We propose to leverage the large amount of readily available unlabeled videos to transfer the temporal information from video domain to static image domain and hence improve static image action recognition. Specifically, We propose a video prediction model to predict the future video of a static image and use the future predicted video to improve static image action recognition.

Honors & Awards

**Full scholarship to pursue M.S. and Ph.D. degrees in the USA**
Kuwait University Scholarship Program  
Khaldiya, Kuwait  
August 2016

**Valedictorian of the Graduates of Excellence Class of 2013/2014**
Graduates of Excellence Ceremony Class of 2013/2014 Under the Patronage and Presence of HH the Amir of Kuwait  
Shuwaikh, Kuwait  
March 2015

**Top 10 students graduating from College of Engineering and Petroleum for the academic year of 2013/2014**
Graduates of Excellence Ceremony Class of 2013/2014 Under the Patronage and Presence of HH the Amir of Kuwait  
Shuwaikh, Kuwait  
March 2015

**Dean’s Honor list and Outstanding Students list**
College of Engineering and Petroleum, Kuwait University  
Khaldiya, Kuwait  
2012 - 2014

**Top Outstanding Student Award**
Kuwait University  
Khaldiya, Kuwait  
2012

**Award of Excellence and Creativity from the Amir Sheikh Sabah AlAhmad Al-Jaber Al-Sabah**
Ministry of Education  
Khaldiya, Kuwait  
June 2013
References

Jia-Bin Huang (Academic advisor)
Professor
Department of Electrical and Computer Engineering
Virginia Tech
Blacksburg, VA
✉️ jbhuang@vt.edu