MSc(CS) Dissertation Public Seminar

****

Title: Cross-frame Selection and Compact Attention for Video-to-video Translation

Speaker: Huang Risheng

Date & Time: April 15 2021, Thursday, 11:00am

Zoom Meeting Link: https://hku.zoom.us/j/95488647812

Meeting ID: 954 8864 7812

Password: 976970

Abstract:

This paper proposes a cross-frame adaptive compact attention model for few-shot video-to-video translation. Existing work in this domain only uses features from pixel-wise attentions without considering the correlations among multiple reference images, which leads to heavy computation but limited performance. Therefore, we introduce a novel adaptive compact attention to efficiently extract contextual features jointly from multiple reference images, of which encoded view-dependent and motion-dependent information can significantly benefit the synthesis of realistic videos. Our core idea is to extract compact basis sets from all the reference images as higher-level representations. To further improve the reliability, in the inference phase, we also propose a novel method based on the Delaunay Triangulation algorithm to automatically select the resourceful references according to the input label. We extensively evaluate our method on a large-scale talking-head video dataset and a human dancing dataset; the experimental results show the superior performance of our method for producing photorealistic and temporally consistent videos, and considerable improvements over the state-of-the-art.

About the Speaker:

Huang Risheng is currently a full-time MSc(CS) student of the Department of Computer Science in the University of Hong Kong. His supervisor is Prof. Wenping Wang.

All are welcome!

Tel: 3917-1828 for enquiries

MSc(CS) Dissertation Public Seminar

****

Title: Preserving Details of Objects in Multiple-objects Text to Image Generation

Speaker: Chen Dongning

Date & Time: April 15 2021, Thursday, 11:40am

Zoom Meeting Link: https://hku.zoom.us/j/95488647812

Meeting ID: 954 8864 7812

Password: 976970

Abstract:

Automatic text to image generation aims at generating images conditioned on given text descriptions. There are numerous applications, such as computer-aided design, photo-editing, and automatic illustration generation. Previous works in this area often target generating an image with a single object, like StackGAN and AttnGAN. Recent research extends this area to generate multiple objects in one image, like Obj-GAN.

However, the objects generated by those models are often distorted and missing important details. In this study, an approach is proposed to improve the quality of the generated image, to preserve the details of objects and to generate photo-realistic images in multiple-objects text to image generation. Two key components are proposed in this study to achieve this goal. The first one is to use edges as part of the inputs in the Image Generation stage. The second one is that a Gate Structure is proposed to offer more flexible control over the filters with respect to the text description. The model built from the proposed approach, named GateGAN, is capable of generating a relatively more realistic image given a text description, and is able to preserve some levels of details in the objects generated. The proposed model reaches 27.43 for Inception Score and 25.81 for Fréchet Inception Distance, which is a large margin ahead of the results of the referenced studies.

About the Speaker:

Chen Dongning is currently a full-time MSc(CS) student of the Department of Computer Science in the University of Hong Kong. His supervisor is Dr. Ping Luo.

All are welcome!

Tel: 3917-1828 for enquiries