MSC projects to be supervised by Kin Hong Wong 2020-21, 2020 June 7

Project title 1: Artificial facial expression generation

When people talk, their facial expressions also change accordingly. We propose to use a neural network approach to train an artificial head (graphical or robotic) to express emotion when the system is reading a text message, listening to music or speech. Facial expression is the combined effect of multiple muscle motions on a face. There are many facial expression datasets online “<https://analyticsindiamag.com/10-face-datasets-to-start-facial-recognition-projects/> “. We can train expressions based on these datasets. Then we may use graphics to produce a face that resemble those expressions. Moreover we can produce a mechanical head that use motors to change the artificial face feature positions to makeup the expression. The system is useful in robotics, games and education.

//////////////////////////////////////////////////////////////////////////////////////

Project title 2 : Music emotion classification using neural networks

The online music industry is actively seeking tools that can effectively classify music according to their emotional effects. It can help to recommend suitable music to users according to one’s taste and mood. Since many on-line music recordings have no labels and even names, classification have to be depended on the music signal itself. And there is a huge commercial interest in this research. The proposed method is to use a database and a neural network to train up a system to achieve the goal. Our related previous work can be found at <http://www.cse.cuhk.edu.hk/~khwong/c2018_IWPR2018_LSTM_music_classification.pdf> .

///////////////////////////////////////////////////

Project title 3: Eye track for the people with disability

If you are reading a webpage and encounter some words you don’t know, you may check it up with a dictionary to find out the meanings. We propose to use eye tracking to locate the word one is gazing and use Optical Character Recognition (OCR) to obtain its spelling and checkup the meaning online. Then, the system can give the verbal description of the meaning of the words through a speaker if the user blinks his/her eye. This research requires the technologies of eye tracking [1] and OCR [3]. It is useful for the people with disability and even general users. References:

[1] <https://pysource.com/2019/01/04/eye-motion-tracking-opencv-with-python/>

[2] <https://www.youtube.com/watch?v=kbdbZFT9NQI>

[3] <https://en.wikipedia.org/wiki/Tesseract_(software)> for OCR

///////////////

Project title 4: Deep learning architectural floor plan images for search and 3D model reconstruction

An architectural floor plan defines the relationships among rooms, spaces, traffic patterns, and other physical features of a building structure. In this project, the student is asked to study various machine learning techniques and devise a novel algorithm to recognize the graphical elements on the floor plan images. It is suggested that the student can start the project based on the source code of the paper   
“Deep Floor Plan Recognition Using a Multi-Task Network with room-Boundary-Guided Attention”   
(<https://arxiv.org/pdf/1908.11025.pdf> ). One direction is to enhance the accuracy of the algorithm as much as possible using the dataset provided. Another direction is to further process the final 3D model to make it more efficient and good enough for virtual reality walkthrough. Hence the use of a head-mount–device (HMD) such as (<https://www.oculus.com/?locale=zh_HK>) may be required. You may also explore how to link up the relations of the recognized floor plan elements and use the extracted elements to build up a floor plan search engine.

////////////////////////

Project title 5: Calibration of multiple Kinect depth sensors for full surface model reconstruction

In this project, we will investigate different methods to calibrate a 3-D scanning system consisting of multiple Azure/Kinect sensors. The main function of the scanning system is for the reconstruction of the full surface (Omnidirectional) model of an object. In this work, we will build a four Azure/Kinect system that the Kinect range sensors are positioned around the target object. Each Kinect is responsible for capturing a small local model, and the local models found will be combined to become the full surface model. Our previous pilot test is described in [1]. It is open for the students to tackle the problem using any methods in machine learning instead of solving the problem in an analytical way. In this sense, learning an artificial intelligence tool like tensor flow is essential. This system will be very useful in robot vision, virtual reality and many other industrial applications.

Reference:

[1]<https://www.cse.cuhk.edu.hk/~khwong/www2/conference/2016/IWPR2016/draft_IWPR2016_cal_mkinects.pdf>