NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

PH.D DEFENCE - PUBLIC SEMINAR

Title: **SUMMARIZATION FROM MULTIPLE USER GENERATED VIDEOS IN GEO-SPACE**

Speaker: Ms Zhang Ying

Date/Time: 18 November 2014, Tuesday, 02:30 PM to 04:00 PM

Venue: Executive Classroom, COM2-04-02

Supervisor : Dr Roger Zimmermann, Associate Professor, School of Computing

Abstract:

In recent years, we have witnessed an overwhelming number of user-generated videos being captured on a daily basis. An essential reason is the rapid development in camera technology and hence videos are easily recorded on multiple portable devices, especially on mobile smartphones. Such flexibility encourages the modern videos to be tagged with additionally various sensor properties. In this thesis, we are interested in geo-referenced videos whose meta-data is closely tied to geographic identifications. These videos have great appeal for prospective travelers and visitors who are unfamiliar with a region or a city. For example, before someone visits a place, a geo-referenced video search engine can quickly retrieve a list of videos that are captured in this place so the visitors could obtain an overall visual impression, conveniently and quickly. However, users face the prospect of an ever increasing viewing burden if the size of these video repositories keeps increasing and as a result more videos are relevant to a search query. To provide viewers with an efficient way to browse these video retrievals, we introduce a novel solution to automatically generate a summarization from multiple user generated videos and present their salience to viewers in an enjoyable manner.

This thesis consists of three major parts. In the first part, we introduce three pieces of work to produce a preview video to summarize a sub-area in geo-space from multiple videos. Several metrics are proposed to evaluate the summary quality and a heuristic method is used to determine the video (segment) selection and connection. One of the key features of our technique is that it leverages the geographic contexts to create a satisfactory summarization result automatically, robustly and efficiently. We also propose a graph based model to formulate this summary problem which can be applied to general videos. In the second part, an interactive and dynamic video exploration system is built where people can conduct personalized summary queries through direct map-based manipulations. In the third part, we investigate whether external crowdsourcing databases contribute to improving the summary quality. Proposing a GMM model and integrating visual or social knowledge, we

recommend a list of locations to be preferentially selected in a summarization as they are of utmost potentials to capture appealing photos.