ABSTRACT

Nowadays people produce a huge number of videos; many are uploaded to the Internet on social media sites such as YouTube and Facebook. There is a strong need to develop automatic solutions for recognizing the contents of these videos. Potential applications of such techniques include effective video content management and retrieval, open-source intelligence analysis, etc. In this talk, I will introduce our recent works on high-level video event recognition. I will start by briefly describing a system that achieved the best performance in the 2010 multimedia event detection task of NIST TREC video retrieval evaluation (TRECVID). After that I will introduce a newly constructed Internet consumer video dataset, on which we try to measure human recognition performance of video events and compare this with the automatic solutions. Finally I will discuss the speed efficiency of several popular techniques and suggest component-level options for “real-time” recognition.

BIOGRAPHY

Yu-Gang Jiang is an associate professor of computer science at Fudan University, Shanghai, China. He finished a PhD in Computer Science at City University of Hong Kong. Before Fudan, he was a postdoctoral research scientist at Columbia University. He was a visiting researcher at University of Amsterdam in summer 2011. His current research is focused on visual/audio media analysis and retrieval. His work has led to a best demo award from ACM Hong Kong, the second prize of ACM Multimedia Grand Challenge 2011, and a recognition by IBM Watson Research as one of ten "emerging leaders in multimedia" in 2009. He has designed a few best-performing video analytic systems at the annual U.S. NIST TRECVID evaluation.