EE368 Project Proposal

Mobile slide-to-video search in ClassX database

Mengyuan Yan
SUID:05949866
Email:mengyuan@stanford.edu

Stanford Center for Professional Development (SCPD) has developed a great database of engineering courses' lecture videos. These videos are very useful for students both on and off campus. Yet in this database it is hard to locate the time stamp where professor was talking about a particular concept, equation, graph, etc. This feature would be very helpful to students reviewing lecture slides. The goal of this project is to enable slide-to-video search on mobile devices based on the SCPD ClassX database. Given a picture of particular slide taken by a mobile device, we return to users the lecture video and time stamp when the slide is shown on screen.

I now have access to lecture videos and corresponding slides for six courses from SCPD. The videos are cut into chunks ranging from 5-20 minutes long. Some of the slides have been annotated with the name of matching video and time stamp via pairwise matching between slides and video keyframes. However this method is extremely slow and not acceptable for mobile applications. Residual enhanced visual vector(REVV)[1] is a fast and memory efficient way of mobile visual search. I want to first implement this method for my problem. Based on preliminary experiments, REVV's retrieval accuracy is not very satisfactory for our database. This could be explained by REVV's incapability of capturing the spatial distribution of features, which is crucial for text images' search. Huizhong Chen and Sam Tsai has discovered new ways to detect and represent text in images with Maximally Stable Extremal Regions[2] and WORD-HOGS[3]. I want to also use this method and compare the retrieval result with those obtained by REVV.

Reference:
[1]David Chen, Sam Tsai, Vijay Chandrasekhar, Gabriel Takacs, Ramakrishna Vedantham, Radek Grzeszczuk, Bernd Girod. "Residual enhanced visual vector as a compact signature for mobile visual search". *Signal Processing Vol.93, Issue 8*, 2013.

[2]Huizhong Chen, Sam Tsai, Georg Schroth, David Chen, Radek Grzeszczuk, and Bernd Girod, "Robust Text Detection in Natural Images with Edge-enhanced Maximally Stable Extremal Regions", *Proc. IEEE International Conference on Image Processing, ICIP-11*, 2011.

[3]Sam Tsai, Huizhong Chen, David Chen, and Bernd Girod, "WORD-HOGS: Word histogram of oriented gradients for mobile visual search", *IEEE International Conference on Image Processing (ICIP)*, 2014.