Aalto University

Multimodal Information Retrieval Approach for Multimedia Resources

Title of the dissertation	Multimodal Concept Detection and Annotation in Image and Video Collections
Contents of the dissertation	The World Wide Web has become a common place for finding information for all kinds of purposes. The amount of data which one user can be dealing with has become large and its size is continuously growing. The relevant data for users have not only become large, but also diverse. Hence, searching relevant information from such large and diverse resources is a critical task. However, users cannot always formulate appropriate queries for finding the desired resources. In order to retrieve relevant information, the semantic relationships of the information in different modalities would need to be known and specified. This thesis approaches the multimodal cross-domain semantic retrieval and fusion problem from the point of view of content-based visual analysis and statistical natural language analysis. It also aims at using cross-domain textual semantics to generate pseudo tags for images to improve the performance of the information retrieval task. The focus of the thesis is in bridging the semantic gap between textual and visual content domains. In our experiments, if the query for image domain is missing or not appropriate, the approach is just the same as ordinal text search. Additionally, the image contents and its textual description do not always match. In order to improve the multimodal information retrieval, the method of pseudo tag generation is proposed. The generation of pseudo tags is based on a text-image semantic map, which is calculated by the co-occurrence of latent topics in text and visual concepts in text-image data. In the experiments, the multimodal information retrieval results were considerably improved by using the pseudo tags.
Field of the dissertation	Computer Science, Pattern Recognition
Doctoral candidate	Satoru Ishikawa, M.Sc. (Tech.)
Time of the defence	14.8.2020 at 12 noon
Place of the defence	Public defense will be organized via Zoom: https://aalto.zoom.us/j/66829855979
Opponent	Professor Joni Kämäräinen, Tampere University, Finland
Custos	Professor Samuel Kaski, Aalto University School of Science, Department of Computer Science
Electronic dissertation	http://urn.fi/URN:ISBN:978-952-60-3954-1
Doctoral candidate's contact information	Satoru Ishikawa, satoru.ishikawa@aalto.fi