Multimedia representation and retrieval using graph-based models

Multimedia representation techniques will be reviewed, implemented and evaluated in the context of multimedia retrieval. The multimedia documents considered will consist of textual and visual data. In specific, the Bag-of-Words (BoW) model will be used for the representation of textual information and the visual-analogue of the BoW model, namely the Bag-of-Visual-Words (BoV), will be applied in the context of content-based image and video representation and retrieval (Sivic and Zisserman, 2003; 2009). The BoV clusters the visual features, such as SIFT descriptors (Lowe, 2004) and creates a visual vocabulary. Recently, the Graph-of-Words (GoW) model (Rousseau and Vazirgiannis, 2013), which is alternative to the BoW model, has shown great efficiency in text retrieval and keyword extraction (Rousseau and Vazirgiannis, 2015). The GoW model will be implemented, keywords will be extracted (using graph k-cores) and the graph similarity between two text documents will be computed. The candidate will also examine the visual analogue of the GoW model. The textual and visual graph-based representation will be fused into the multimodal graph-based representation of text-image news items.

**References**

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, *60*(2), 91-110.

Rousseau, F., & Vazirgiannis, M. (2013, October). Graph-of-word and TW-IDF: new approach to ad hoc IR. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management* (pp. 59-68). ACM.

Rousseau, F., & Vazirgiannis, M. (2015). Main Core Retention on Graph-of-Words for Single-Document Keyword Extraction. In *Advances in Information Retrieval* (pp. 382-393). Springer International Publishing.

Sivic, J., & Zisserman, A. (2003, October). Video Google: A text retrieval approach to object matching in videos. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on* (pp. 1470-1477). IEEE.

Sivic, J., & Zisserman, A. (2009). Efficient visual search of videos cast as text retrieval. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *31*(4), 591-606.