Προτεινόμενα Θέματα Διπλωματικής Εργασίας

Από

[Θεόδωρο Γιαννακόπουλο, Ερευνητή ΕΚΕΦΕ Δημόκριτος]

A. Examination of cross-dataset generalization abilities for Speech Emotion Recognition systems

Abstract: Speech Emotion Recognition is associated with predicting the underlying emotion levels that are contained in an input speech signal. There is a range of datasets for this task that are based on different design choices. More specifically, datasets differ, along with other aspects, with respect to (i) the type of speech (acted, improvised, etc), (ii) the definitions of emotional levels, (iii) the context (e.g., theater actors are more expressive compared to everyday situations), (iv) the spoken language, and (v) the recording setup. The scope of this thesis is to empirically examine the cross-dataset generalization ability of models trained under different dataset types and investigate parameters that affect the transfer of knowledge among speech emotion recognition datasets.

B. Automatic, content-based sound retrieval for visual artworks

Abstract: The goal of this thesis is to train and evaluate deep learning models that are able to identify visual attributes from artwork and map them to matching sounds that are "relevant" in terms of content, style, mood and other contextual characteristics. Towards this end, we will consider using existing datasets but we will also build new ones, particularly focusing on annotating explicit correlations between visual and auditory characteristics. This will enable us to train models that provide a multimodal embedding space for both visual and auditory information of artworks. The thesis will be jointly supervised by Foteini Salmouka (f.salmouka@central.demokritos.gr).