Extracting and integrating multimodality features via multidimensional approach for video retrieval

ABSTRACT

This work discusses the application of an Artificial Intelligence technique called data extraction and a process-based ontology in constructing experimental qualitative models for video retrieval and detection. We present a framework architecture that uses multimodality features as the knowledge representation scheme to model the behaviors of a number of human actions in the video scenes. The main focus of this paper placed on the design of two main components (model classifier and inference engine) for a tool abbreviated as VSAD (Video Action Scene Detector) for retrieving and detecting human actions from video scenes. The discussion starts by presenting the workflow of the retrieving and detection process and the automated model classifier construction logic. We then move on to demonstrate how the constructed classifiers can be used with multimodality features for detecting human actions. Finally, behavioral explanation manifestation is discussed. The simulator is implemented in bilingual; Math Lab and C++ are at the backend supplying data and theories while Java handles all front-end GUI and action pattern updating.