Abstract

In recent years, the multimedia retrieval community is gradually shifting its emphasis from analyzing one media source at a time to exploring the opportunities of combining diverse knowledge sources from correlated media types and context. In order to combine multimedia knowledge sources, two basic issues must be addressed: *what* to combine and *how* to combine. While considerable effort has been expended to generate a wide range of ranking features from knowledge sources, relatively less attention has been given to the problem of finding a suitable strategy to combine them. It has always been a significant challenge to develop principled combination approaches and capture useful factors such as query information and context information in the retrieval process.

This thesis presents a conditional probabilistic retrieval model as a principled framework to combine diverse knowledge sources. This model can integrate multiple forms of ranking features (query dependent and query independent features) as well as query information and context information in a unified framework with a solid probabilistic foundation. Under this retrieval framework, we overview and develop a number of state-of-the-art approaches for extracting ranking features from multimedia knowledge sources. In order to deal with heterogenous features, a discriminative learning approach is suggested for estimating the combination parameters. Moreover, an efficient rank learning approach has been developed to explicitly model the ranking relations in the learning process with much less training time.

To incorporate query information in the combination model, this thesis develops a number of *query analysis* models that can automatically discover mixing structure of the query space based on previous retrieval results, and predict combination parameters for unseen queries. In more detail, we propose the *query-class based analysis* model which needs to manually define the query classes and a series of *probabilistic latent query analysis*(pLQA) models which can automatically discover latent query classes from the development data by unifying the combination weight optimization and query class categorization into a discriminative learning framework. To adapt the combination function on a per query basis, this thesis also presents a *probabilistic local context analysis*(pLCA) model to automatically leverage additional retrieval sources to improve initial retrieval outputs. A pLCA variant is proposed to utilize human feedback to adjust combination parameters.

All the proposed approaches are evaluated on multimedia retrieval tasks with large-scale video collections. Beyond multimedia collections, we also evaluate our approaches on meta-search tasks with large-scale text collections. Experimental evaluations demonstrate the promising performance of the probabilistic retrieval framework with query analysis and context analysis in the task of knowledge source combination. The applicability of the proposed methods can be extended to many other areas, such as question answering, web IR, cross-lingual IR, multi-sensor fusion, human tracking, and so forth.