University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Luka Popovic

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, February 15, 2019

Master's Basic Module
Datenbanktechnologie

Topic: Measuring the Similarity of Movies Based on Multidimensional Features

When recommending movies to users in systems such as MovieLens or IMDB, the movies are chosen either according to previous choices of (other) users or based on textual descriptions or tags. Consequently, the systems rely on human-generated information rather than the actual content of the movies.

The goal of this Master's basic module is to investigate alternatives to the existing approaches, more concretely to look at content-based similarity measures for movies using various visual features of key frames or scenes in the movies and also to consider the order in which these features appear.

The work is structured into the following tasks with corresponding outcomes and deliverables:

- **T1: Analyzing existing data**
  This project does not start from scratch but builds on top of some work done by a BSc student at the Free University of Bozen-Bolzano. The first step is to analyze the existing data, to have a closer look at the extracted features, and to determine the suitability of these features for similarity
  Outcome: description and evaluation of the existing features
- **T2: Create new feature set**
  Based on the evaluation in T1, the task here is to decide which features to keep, which features to discard, which features to replace, and which features to add.
  Outcome: a new set of features for measuring similarity with a brief justification for each

- **T3: Feasibility study/experimental evaluation**
  After the first two steps, the goal of the third step is to test the suitability of the new feature set with respect to their suitability for a similarity measure. This may involve more than one iteration between T2 and T3.
  Outcome: rank the features in the new feature set according to their suitability

Supervisor: Sven Helmer

Start date: 18.02.2019

End date: 17.05.2019

Exam: 30.04.2019 15:00-16:00 BIN 2.E.13

Department of Informatics, University of Zurich

Prof. Dr. Michael Böhlen