IMOTION

Heiko Schuldt, University of Basel, Switzerland
heiko.schuldt@unibas.ch
IMOTION at a Glance

• Project Title
Intelligent Multimodal Augmented Video Motion Retrieval System (IMOTION)

• Project Start and Duration
January 1st, 2014 – December 31st, 2016 (3 years)

• Project Partners
  – University of Basel (UNIBAS), Switzerland (Coordinator)
  – Koç University (Koç), Istanbul, Turkey
  – University of Mons (UMONS), Belgium
Scientific Background

• Large Multimedia Collections
  – Private use, professional applications, education

• Content-based Multimedia Retrieval
  – Manual annotations not feasible
  – Use of inherent features (colors, shapes, objects, etc.)
  – Strong focus on image features

• Video Retrieval
  – **Motion** distinguishes video from still images
  – Efficient and effective content-based retrieval of (parts of) videos based on **motion specification** is lacking

→ IMOTION will develop and evaluate
  **Innovative Multi-Modal User Interfaces for Interacting with Videos**
Key Challenges & Potential Impact

Goal: Support novel types of motion queries in video collections

- **User Interaction for Query Specification**
  - Users can specify motion paths of objects in videos
  - via sketches, gestures, natural language, or combinations

- **Machine Learning for Motion Features**
  - Extraction of high-level motion descriptors
  - based on DNN (ConvNets, dimensionality reduction)

- **Information Retrieval & Data Management for Efficient Search**
  - Index structures to jointly support video features and motion metadata
  - Distributed IR Engine: **Scalability** to very large collections

- **Quantitative and Qualitative Evaluation**
Competences needed for IMOTION:

- **User Interaction**
  - Sketch-based Interfaces
  - Speech-based Interfaces

- **Information Retrieval & Databases**
  - Indexing, Retrieval Models
  - Distributed Data Management

- **Machine Learning**
  - Feature Extraction
  - Concept classification
Databases and Information Systems Group (Prof. Heiko Schuldt)

- **Multimedia Retrieval**, especially for **Big Data Collections**
  - *Video Similarity Search* [RGS 14, RGH+ 16]
  - *Query Types*: Seamless Combination of Boolean Retrieval and Similarity Search [GAS 14b]
  - *Very Large Collections*: Distribution in Map/Reduce Style [GAS 14a], progressive queries [GS 16]
  - *Dealing with ambiguity* [RTS 16]
  - *Video Feature Extraction* [SGS 14]
  - *Video Feature Combination* [RGS 14]

- **User Interfaces for Multimedia Queries**
  - Sketches: Interactive Paper, Tablets, Mobile Devices
  - Gestures: IR-based gesture recognition
Intelligent User Interfaces Laboratory (Prof. T. Metin Sezgin)

- **User Interfaces for Sketch-based Interaction**
  - Sketch Recognition
  - Early recognition through auto-completion

- **Multimodal Motion Retrieval**
  - *Sketch Interfaces*: Retrieval of 3D articulated shapes through 2D sketch-based input.
  - *Speech Interfaces*: Description of motion queries
  - *Multimodal fusion*: Co-reference resolution.
  - *Use case*: Soccer videos
University of Mons

Research Center for Creative Technologies (Dr. Stéphane Dupont)

- **Deep Neural Networks (DNNs) for Multimedia Retrieval**
  - Feature extraction (often unsupervised learning): learning representations of sketch/image/video content
  - Classification (supervised learning): recognizing concepts and actions
  - Regression (supervised learning): mapping sketch queries to real content features

- **Latest Developments**
  - *Improved sketch recognition [SDM 15, SDM 16b]*
    - using temporal nature of sketching process
    - comparative study of fusion schemes
    - sketch-to-image search using common ConvNet features
  - *Improved action recognition [SKP+ 15, SDM 16a]*
    - comparative study of optical flow algorithms
    - fusion of spatial and temporal streams
  - *Beyond single concept: multi-label multi-instance*
    - new approach & benchmarking under development

Inspired by [Le et al. 2012]
... Consortium as a Whole

- Small, but highly complementary consortium
IMOTION Workplan

WP 2 (UMONS)
- Feature Extraction

WPs 3 & 4 (Koç)
- UI for Query Specification

WPs 5 & 6 (UNIBAS)
- Query Execution

WP 7 (all)
- Build and continuously update integrated system

WP 8 (all)
- Evaluate integrated system
**IMOTION System: Query Types**

- **Query-by-Sketch** using a rough hand-drawn sketch

- **Query-by-Sketch** with **Query-by-Example** by adding/removing parts from sample image

- **Query-by-Example** based on query image or a result object

- **Query-by-Motion** motion of objects across frames via flow fields

**Relevance Feedback** refining query results by marking relevant and irrelevant elements from the result list
IMOTION System: Overview

Online
- Query Sketch / Image
- Ranked Result List
- Video Decoder
- Video Segmenter
- Extraction Runtime
- Retrieval Runtime
- Query Generation
- Score Fusion and Re-ranking
- Features
  - Semantic Image Features
  - Motion Features
  - Color / Edge Features
  - ...

Offline
- Semantic Image Features
- Motion Features
- Color / Edge Features
-...

SQL Query
- Results
- Ranked NN List

Feature Vector Database

Features

Ranked NN List

SQL Query
- Results

Main Results & Achievements

• 15 scientific publications

• Participation at the 4th Video Search Showcase (VSS) competition [RGS+15], co-located with the Multimedia Modeling Conference (MMM’15)
  – IMOTION finished in second place overall and has been winner of the visual search task

• Participation at the 2016 Video Browser Showdown (VBS), co-located with MMM’16
  – IMOTION has participated in two modes: manually operated [RGH+ 16] and fully automated [RGT+ 16]

• Best demo award at MMM’16 [RGH+ 16]

• Demo at IUI conference on sketch auto-completion [TGR+ 16]

• Open source version of search engine and database backend (vitrivr) supported by Google Summer of Code 2016

• Plan: evaluation in TRECVID 2016, VBS 2017
Project Management ...

- Project has started on January 1st, 2014

- Staffing
  - UNIBAS
    - 1 PostDoc for 2 years (C. Tănase)
    - 2 PhD students for 3 years (I. Giangreco, L. Rossetto)
  - Koç
    - 1 PostDoc for 2 years (Y. Sahillioğlu)
    - 1 MSc student (O. Altıok)
  - UMONS
    - 1 PhD student for 3 years (O. Seddati)

- Internal project meetings (twice a year)
  - February 2014 kick-off in Basel
  - 09/2014 in Mons, 01/2015 in Istanbul, 08/2015 in Mons, 04/2016 in Basel
  - plus bilateral technical meetings
... Project Management

- **Student/scientists exchanges**
  - eENTERFACES 2015 in Mons (August 2015): four week on-site collaboration

- **Financial reporting**
  - UNIBAS and UMONS: in plan
  - Koç: underspent

- **Dissemination, Infrastructure, Collaboration**
  - Project Website: imotion-project.eu
  - Repository, mailing list, etc. has been set up

- **Sustainability**
  - Project deliverables published as scholarly papers
  - IMOTION software in open source license
  - Collections and evaluation metrics released to the research community
    OSVC: Open Short Video Collection
IMOTION Publications (1/4)


IMOTION Publications (3/4)


IMOTION Publications (4/4)


