

# 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 1<sup>st</sup>, 2014 – December 31<sup>st</sup>, 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 **motion ontology**
- **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**
  - **User studies** based on collections released by the project consortium



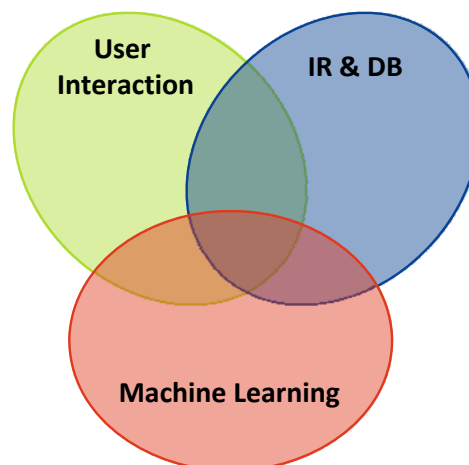
„Show me all shots on goal from the penalty area, with support from the left wing“

## Consortium as a Whole ...



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



## University of Basel



Databases and Information Systems Group (Prof. Heiko Schuldt)

- **Multimedia Retrieval, especially for Big Data Collections**
  - *Query Types*: Seamless Combination of Boolean Retrieval and Similarity Search [GAS 14b]
  - *Very Large Collections*: Distribution in Map/Reduce Style [GAS 14a]
  - *Video Feature Extraction* [SGS 14]
  - *Video Feature Combination* [RGS 14]
  - Cloud Data Management: Microsoft Azure Research Grant
- **User Interfaces for Multimedia Queries**
  - Sketches: Interactive Paper, Tablets, Mobile Devices
  - Gestures: IR-based gesture recognition



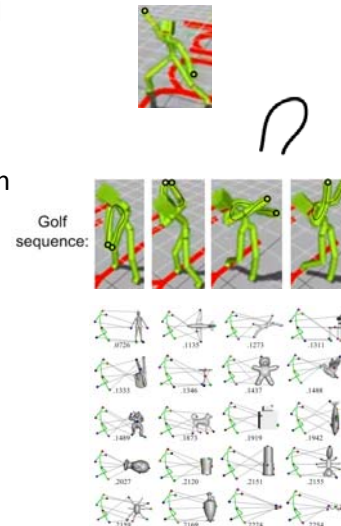
## Koç University



Intelligent User Interfaces Laboratory (Prof. T. Metin Sezgin)

- **User Interfaces for Sketch-based Motion Retrieval**

- Identify discriminative points in a mesh representation of an object
- *Query*: Sketch of a curve that represents a path between the feature points during desired action
- *Retrieval*: select sequences in which this path occurs. Comparison between sketch and DB objects



CHIST-ERA Projects Seminar 2015, Madrid — IMOTION — March 19, 2015

7

## University of Mons



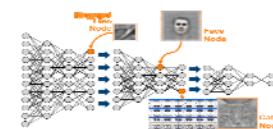
Research Center for Creative Technologies (Dr. Stéphane Dupont)

- **Deep Neural Networks (DNNs)**

- Feature extraction / regression / classification schemes to support multimedia information retrieval using motion
- Set-up of workflow to enable DNNs (*ConvNets*) on video

- **Machine Learning for Multimedia Retrieval**

- *Unsupervised* learning (feature extraction): learning representations of content
- *Supervised* learning (classification): recognize concepts and actions
- *Supervised* learning (regression): map sketch queries with real content features



Inspired by [Le et al. 2012]



- **Large-scale Video Machine Learning**

- 5D data
- *DeconvNets auto-encoders* for video

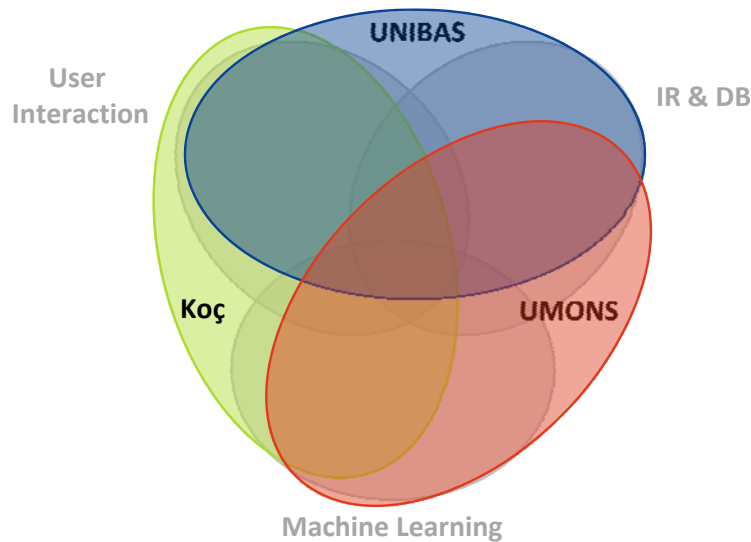
CHIST-ERA Projects Seminar 2015, Madrid — IMOTION — March 19, 2015

8

## ... Consortium as a Whole



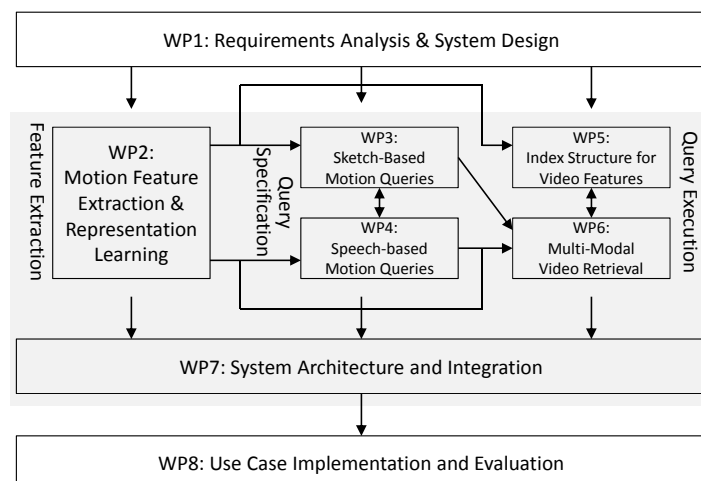
- Small, but highly complementary consortium

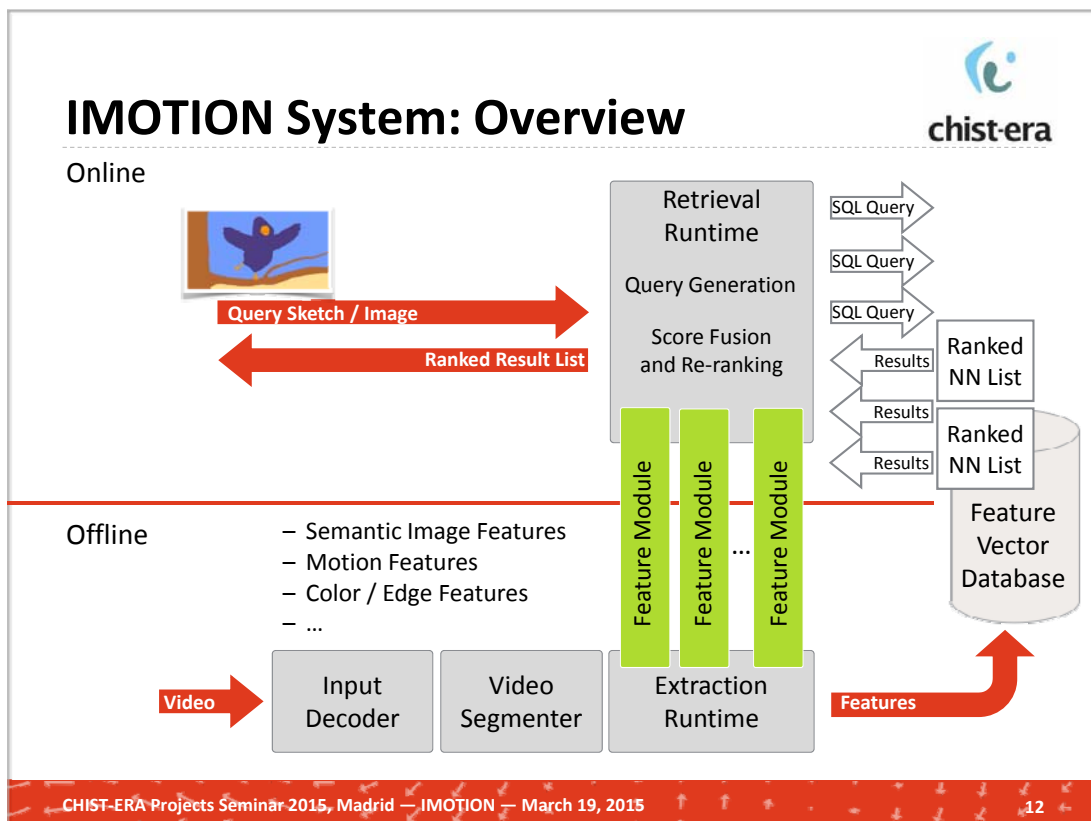
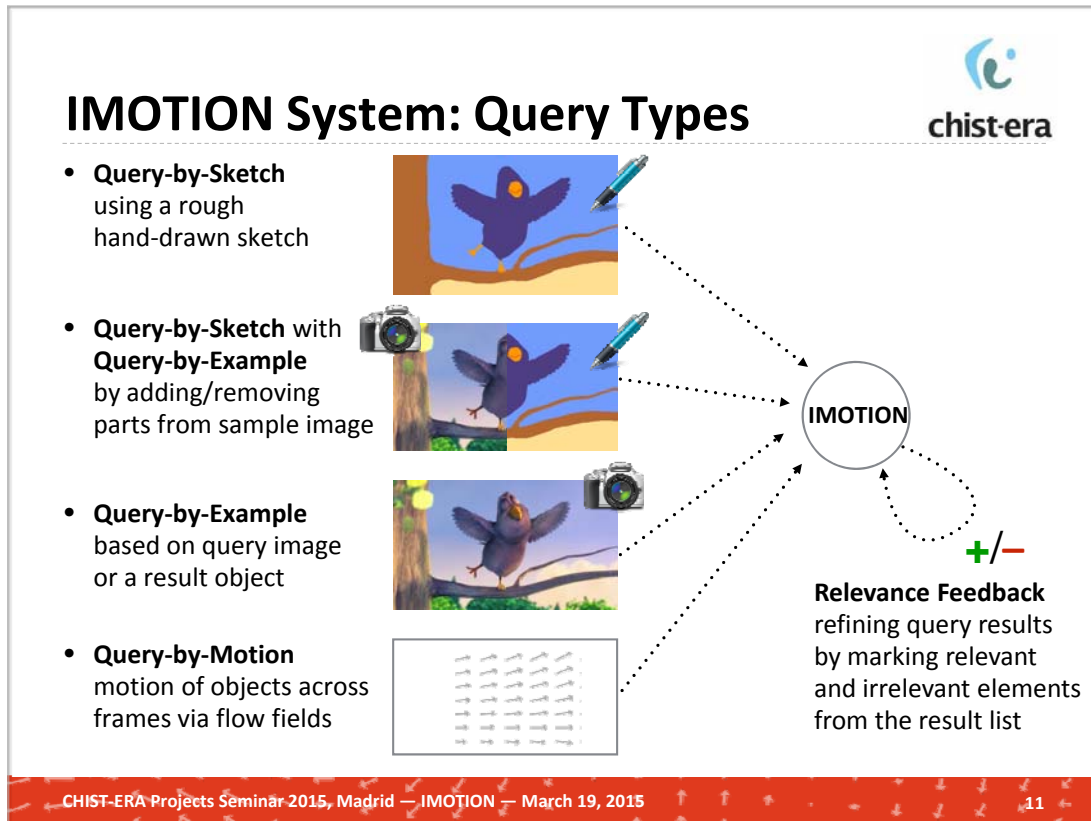


## IMOTION Workplan



- ✓ WP 1 (all)
  - Requirements
- 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

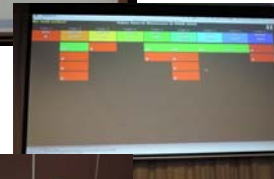




## IMOTION@VSS 2015



- Participation at the 4<sup>th</sup> Video Search Showcase competition [RGS+15]
  - Part of the Multimedia Modeling Conference (MMM)
  - Two types of search tasks
    - Visual Known Item Search (sketches and/or captures)
    - Textual Search (to be addressed via sketches)
    - Both tasks performed by experts and by novices
  - IMOTION finished in second place overall (1'213 points vs. 1'223 points), and has been clear winner of the visual search task



## Project Management ...



- **Project has started on January 1<sup>st</sup>, 2014**
- **Staffing**
  - UNIBAS
    - 1 PostDoc for 2 years (C. Tănase)
    - 2 PhD students for 3 years (I. Giangreco, L. Rossetto)
  - KOÇ (staffing will be completed in next PhD student intake)
    - 1 PostDoc for 2 years (Y. Sahillioğlu)
    - 2 PhD students for 3 years (NN, NN)
  - UMONS
    - 1 PhD student for 3 years (O. Seddati)
- **Internal project meetings**
  - February 2014 kick-off in Basel
  - September 2014 in Mons
  - January 2015 in Istanbul
  - Plus bilateral meetings for the VSS preparation

## ... Project Management



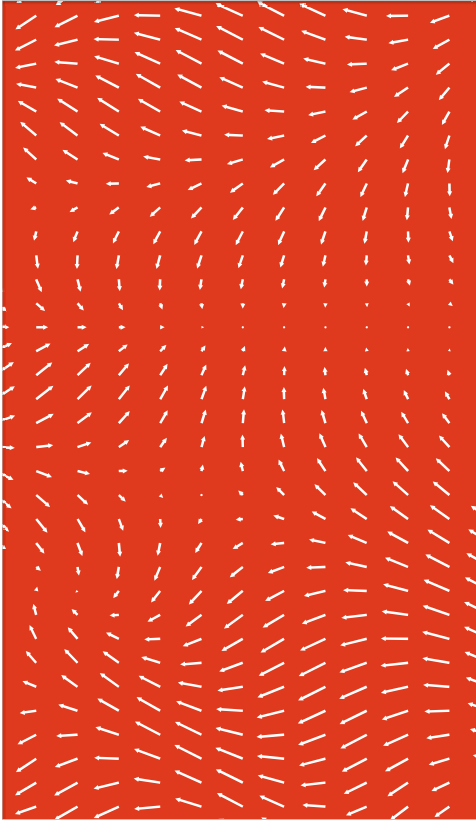
- **Student/scientists exchanges**
  - eINTERFACES 2015 in Mons (August 2015): four week on-site collaboration  
[www.interface.net/enteface15/wp-content/uploads/2015/02/VideoSketcher.pdf](http://www.interface.net/enteface15/wp-content/uploads/2015/02/VideoSketcher.pdf)
- **Financial reporting**
  - Nothing to report (all partners in plan)
- **Dissemination, Infrastructure, Collaboration**
  - Project Website: [imotion-project.eu](http://imotion-project.eu)
  - Repository, mailing list, etc. has been set up
  - Consortium Agreement has been signed
- **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



- [RGS\*15] Luca Rossetto, Ivan Giangreco, Heiko Schuldt, Stéphane Dupont, Omar Seddati, Metin Sezgin, Yusuf Sahillioğlu: *IMOTION – a Content-based Video Retrieval Engine*. In: Proceedings of the 21<sup>st</sup> MultiMedia Modelling Conference (MMM2015) - Video Search Showcase Track, Sydney, Australia, pp. 255-260, January 2015, Springer LNCS, Vol. 8936.
- [GAS 14a] Ivan Giangreco, Ihab Al Kabary, Heiko Schuldt: *ADAM — A Database and Information Retrieval System for Big Multimedia Collections*. In: Proceedings of the 3<sup>rd</sup> International Congress on Big Data, Anchorage, USA, pp. 406 – 413, June 2014.
- [GAS 14b] Ivan Giangreco, Ihab Al Kabary, Heiko Schuldt. *ADAM — A System for Jointly Providing IR and Database Queries in Large-Scale Multimedia Retrieval*. In: Proceedings of the 37<sup>th</sup> International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '14), Gold Coast, Australia, July 2014.
- [RGS 14] Luca Rossetto, Ivan Giangreco, Heiko Schuldt: *Cineast: A Multi-Feature Sketch-Based Video Retrieval Engine*. In: Proceedings of the 16th IEEE International Symposium on Multimedia (ISM2014), Taichung, Taiwan, December 2014. IEEE.
- [SGS 14] Fabio Sulser, Ivan Giangreco, Heiko Schuldt: *Crowd-based Semantic Event Detection and Video Annotation for Sports Videos*. In: Proceedings of the 3<sup>rd</sup> International ACM Workshop on Crowdsourcing for Multimedia, Orlando, USA, pp. 63-68, November 2014.





  
chist-era

THANK YOU



<http://www.shutterstock.com>