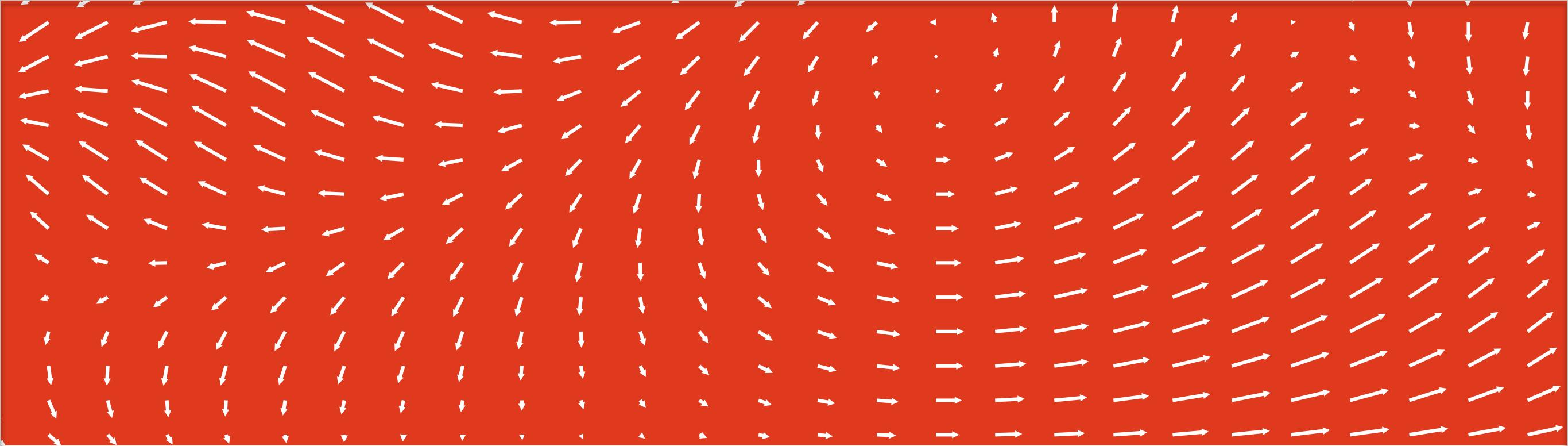




IMOTION

Ivan Giangreco, University of Basel, Switzerland

<ivan.giangreco@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)
extension of IMOTION project until **December 2017**,
extension of Swiss part until **May 2018**

- 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



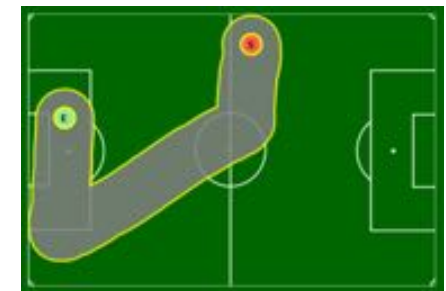
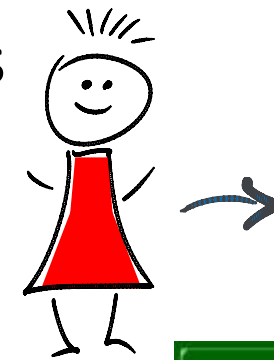
↳ **IMOTION will develop and evaluate**

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**

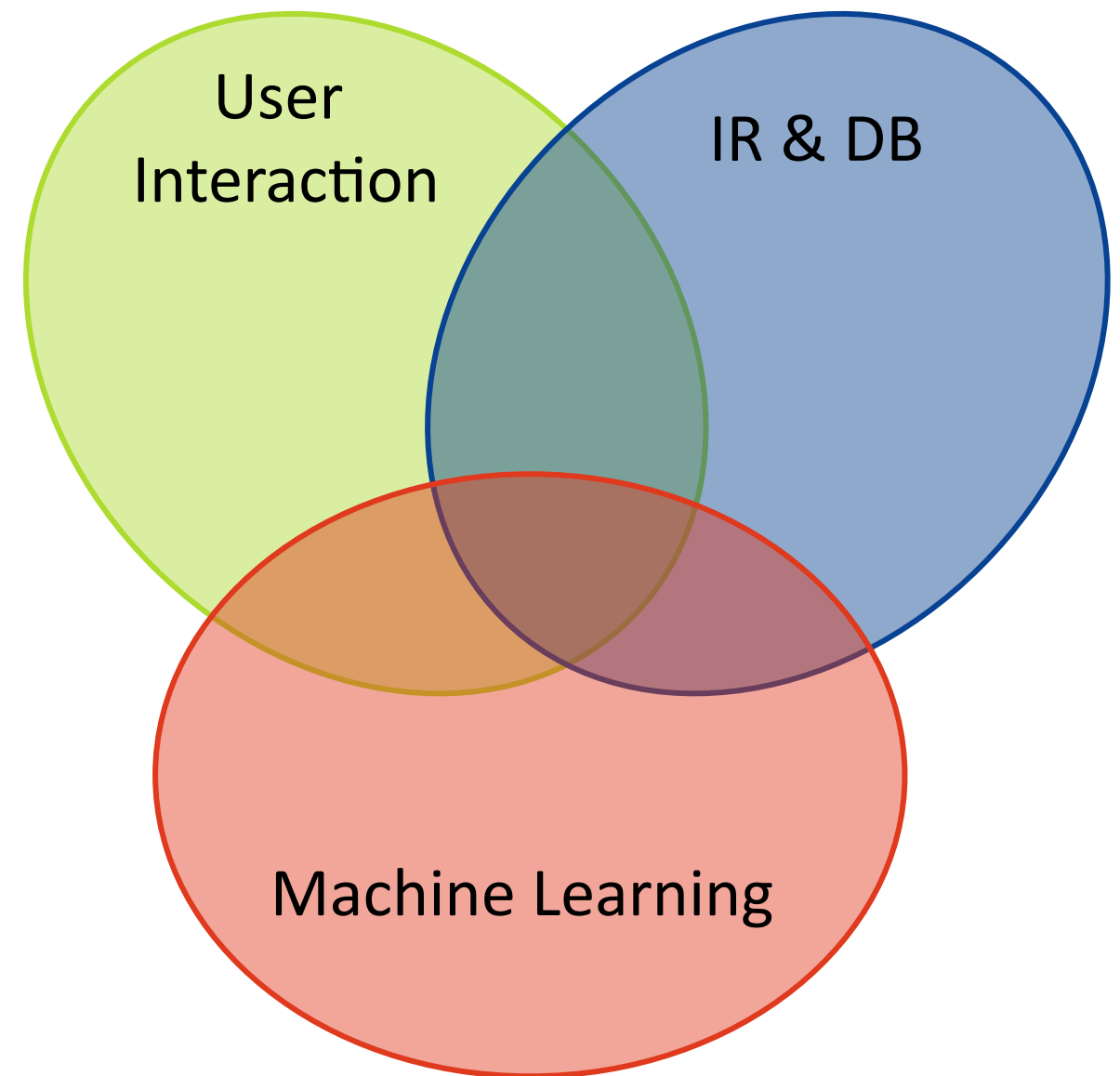


„Show me all shots
on goal from the
penalty area, with
support from the
left wing“

Consortium as a Whole ...

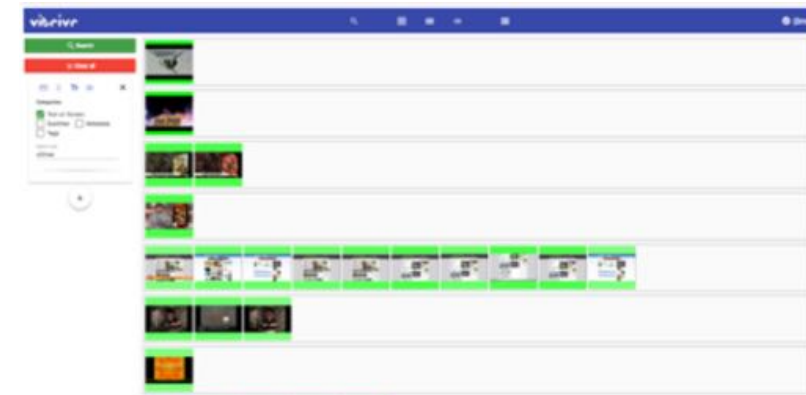
Competences needed for IMOTION:

- **User Interaction**
 - Sketch-based Interfaces
 - Speech-based Interfaces
- **Information Retrieval & Databases**
 - Indexing, Retrieval Models
 - Distributed Data Management
- **Machine Learning**
 - Feature Extraction
 - Concept classification



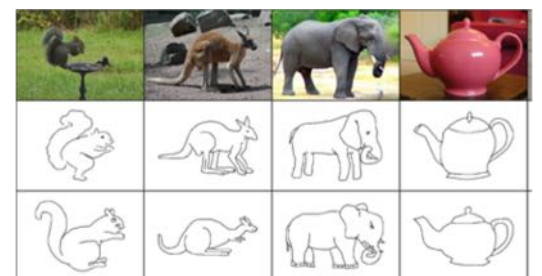
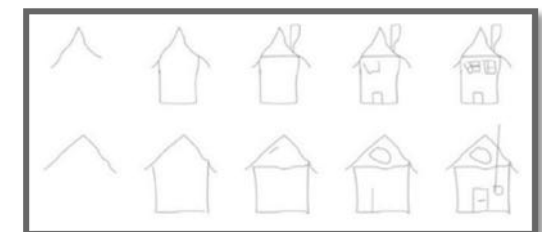
Databases and Information Systems Group (Prof. Heiko Schuldt)

- **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]
 - *Joint Sketch/Speech Interface* [GGR⁺ 17]
 - *Large-Scale Datasets* [RS⁺ 17], [RH 17]
- User Interfaces for Multimedia Queries
 - Sketches: Interactive Paper, Tablets, Mobile Devices
 - Gestures: IR-based gesture recognition



Research Center for Creative Technologies (Dr. Stéphane Dupont)

- **Deep Neural Networks (DNNs) for Multimedia Retrieval**
 - Feature extraction: learning representations of sketch/image/video content
 - Classification: recognizing concepts and actions
 - Regression: mapping sketch queries to real content features
- **Sketch Recognition and Sketch-based Retrieval**
 - *Improved sketch recognition [SDM 17a]*
 - using temporal nature of sketching process
 - comparative study of fusion schemes
 - *Improved Sketch-based image retrieval (SBIR) [SDM 17{a,b,c}]*
 - sketch-to-image search using ConvNet features
 - multi-task quadruplet networks for SBIR
- **Concept and Action Recognition**
 - *Improved action recognition [SDM 16a]*
 - comparative study of optical flow algorithms
 - fusion of spatial and temporal streams
 - *Beyond single concept: multi-label multi-instance recognition and localization*



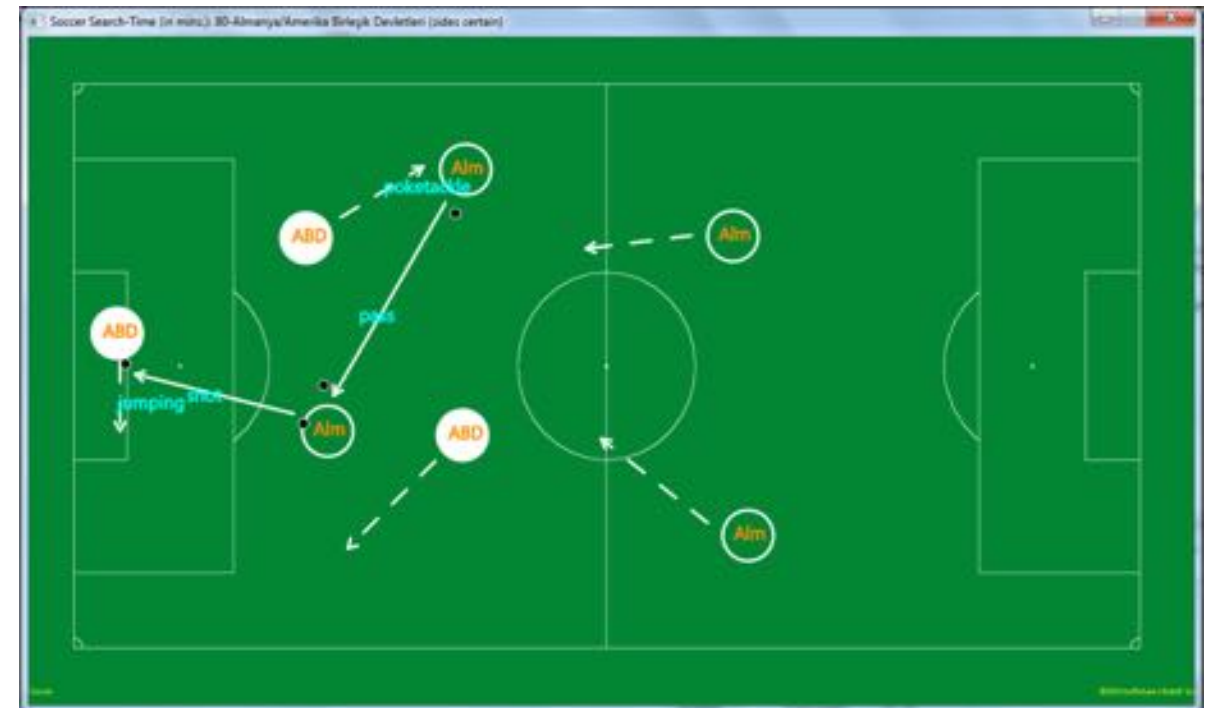
Intelligent User Interfaces Laboratory (Prof. T. Metin Sezgin)

- **User Interfaces for Sketch-based Interaction**

- Early recognition of sketched symbols through auto-completion

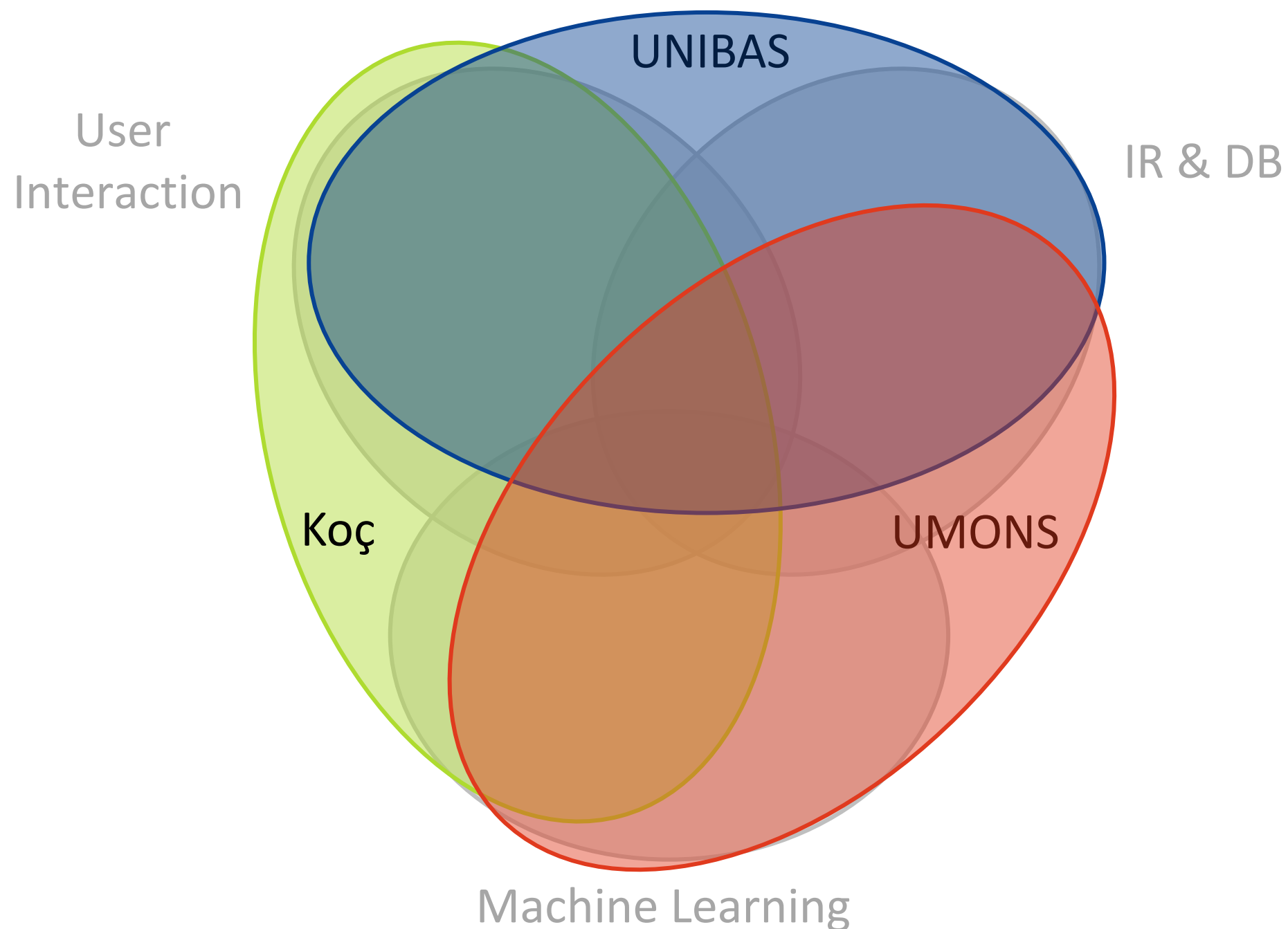
- **Multimodal Motion Retrieval**

- User interface for soccer match retrieval through speech and sketching
 - Sketching: Motion of the ball and players
 - Speech: Body movement of players and details about ball motion (e.g. shots, passes, goal kicks etc.)



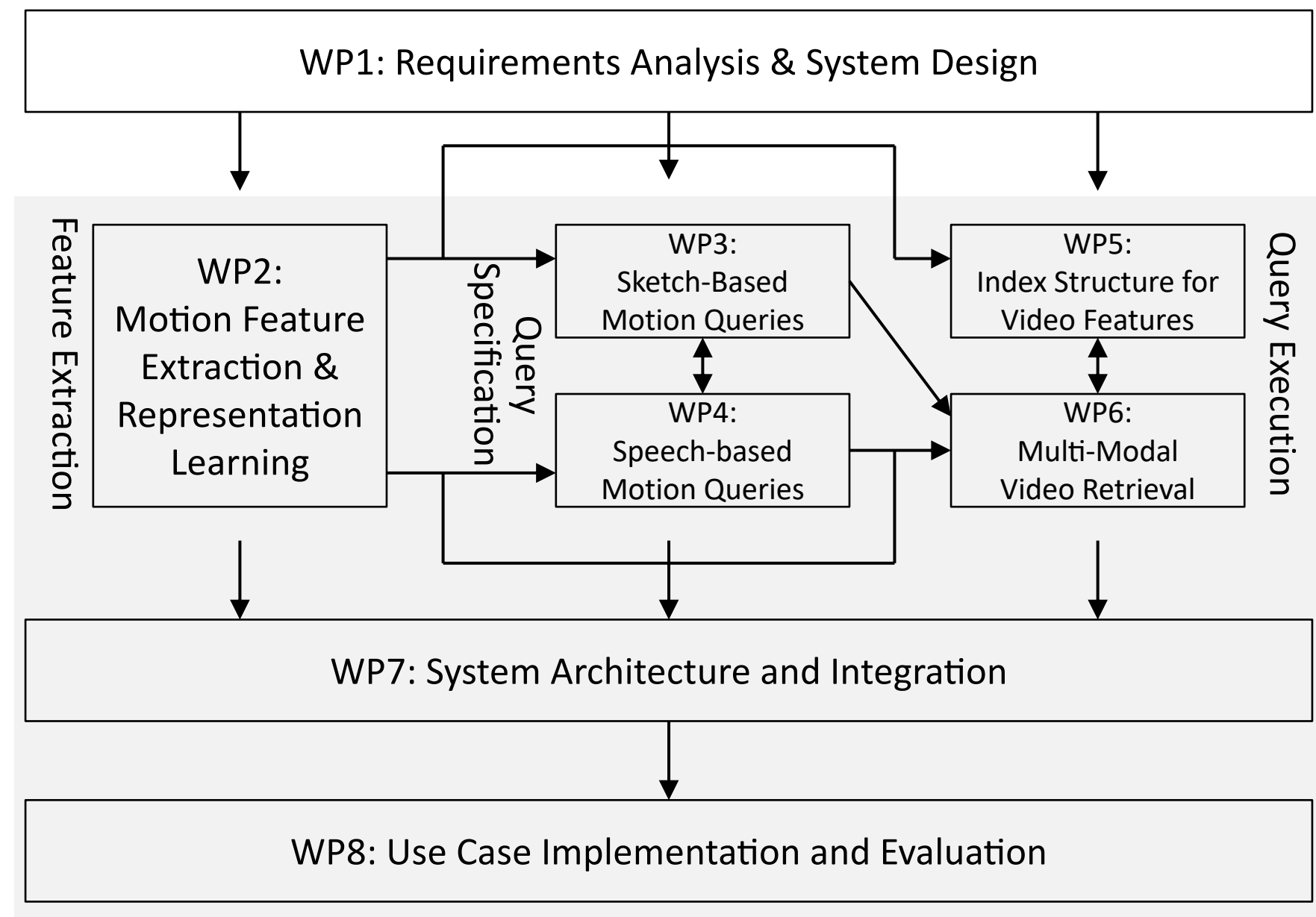
... Consortium as a Whole

- Small, but highly complementary consortium



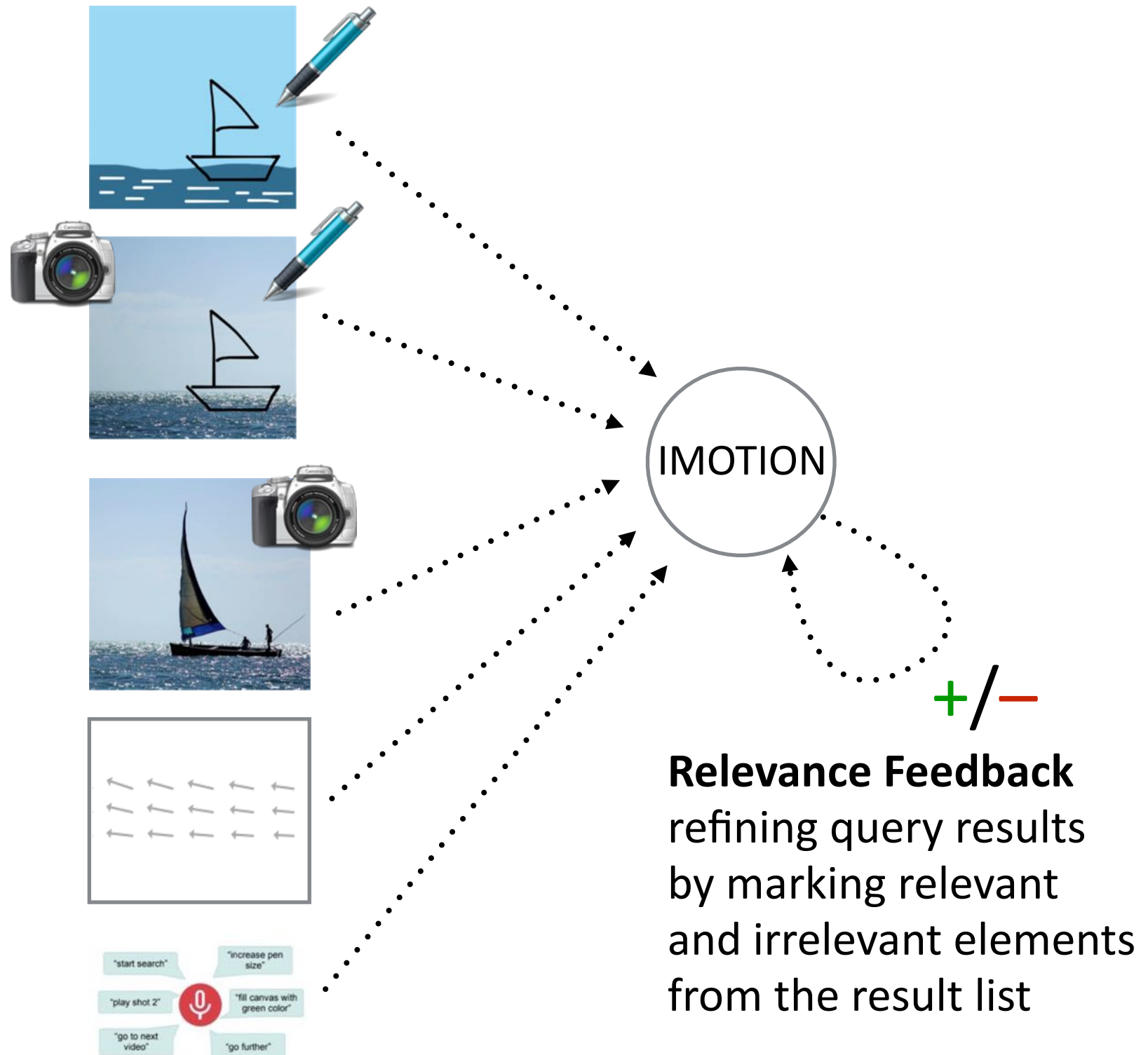
IMOTION Workplan

- 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)
 - Evaluation
- New activities: extension to additional media types



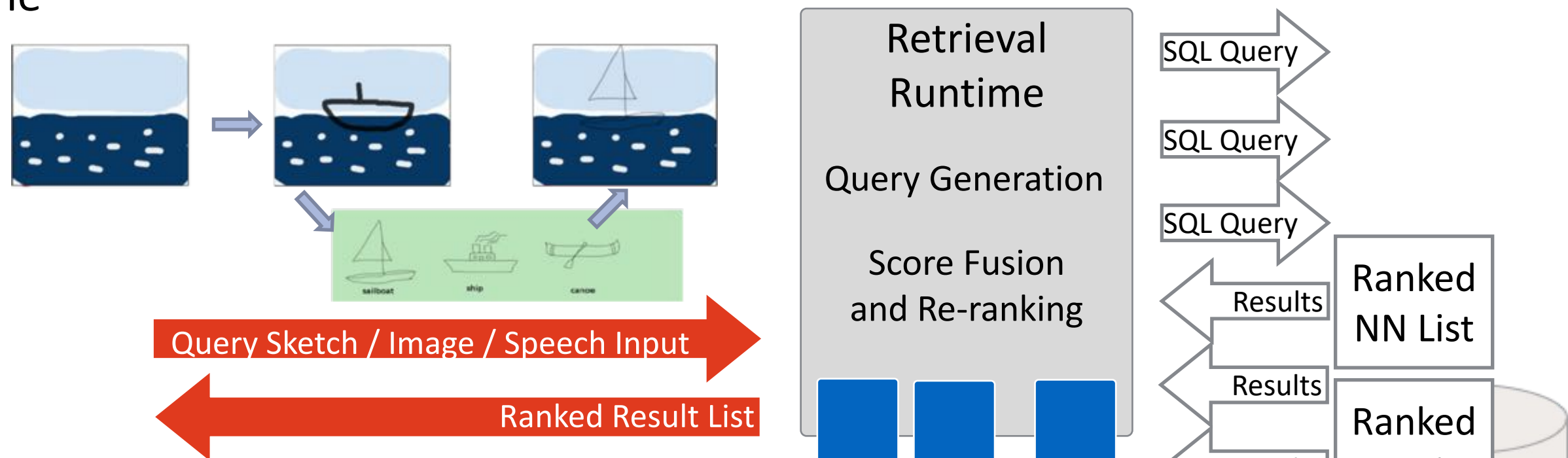
IMOTION System: Query Types

- **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
- **Spoken Queries**
natural language



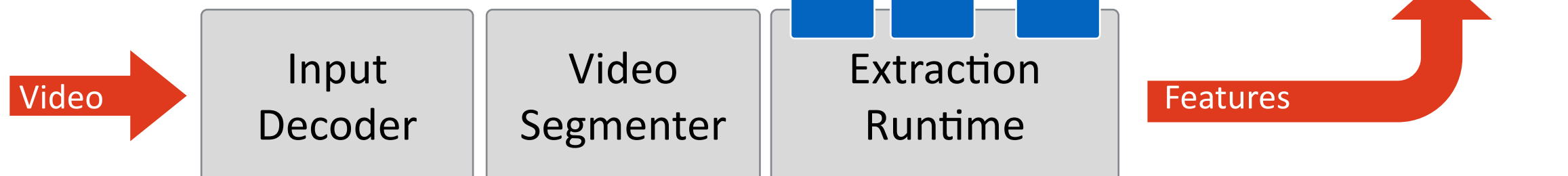
IMOTION System: Overview

Online



Offline

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



Main Results & Achievements

- > 30 scientific publications
- Participation to the 6th Video Browser Showdown [RGR⁺ 18]
- **Winner** of the 2017 5th Visual Browser Showdown [RGT⁺ 17], co-located with the Multimedia Modeling Conf. (MMM'17)
 - TRECVID AVS queries & visual / textual known item search
- Participation at the 4th Video Search Showcase (VSS) competition [RGS⁺15], co-located MMM'15 (2nd place)
- Participation at the 2016 Video Browser Showdown (VBS) (manually [RGH⁺ 16] and fully automated [RGT⁺ 16])
- Participation at TRECVID AVS 2016 [TRG⁺ 16, TRG⁺ 16b]
- **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 (vitivr) [RGT⁺ 16b], supported by Google Summer of Code 2016 [GGR⁺ 17] and 2018



Project Management ...

- **Project has started on January 1st, 2014**
- **Staffing**
 - UNIBAS
 - 1 PostDoc for 2 years (C. Tănase)
 - 2 PhD students for 3+1 years (I. Giangreco, L. Rossetto)
 - Koç
 - 1 PostDoc for 2 years (Y. Sahillioğlu)
 - 2 research students (O. Altıok, A.Tosun)
 - UMONS
 - 1 PhD student for 3 years (O. Seddati)
- **Internal project meetings: twice a year**
 - Plus bilateral technical meetings and Skype conferences
- **Financial reporting**
 - UNIBAS and UMONS: in plan
 - Koç: underspent

... Project Management ...

- **Student/scientists exchanges**
 - eNTERFACE 2015 in Mons (August 2015): four week on-site collaboration
www.enterface.net/enterface15/wp-content/uploads/2015/02/VideoSketcher.pdf
 - Joint PhD thesis (Cotutelle) UMONS & UNIBAS
- **Dissemination, Infrastructure, Collaboration, Awards**
 - Project Website: imotion-project.eu
 - Repository, mailing list, project management tools
 - Parts of IMOTION software (vitivr) in open source license
 - Participation at Google Summer of Code 2016, 2018
 - Participation to Video Browser Showdown 2015, 2016, 2017, 2018
 - **Winner of the Video Browser Showdown (VBS) 2017**
- **Sustainability**
 - Project deliverables published as scholarly papers
 - Collections and evaluation metrics released to the research community
OSVC: Open Short Video Collection; [RS⁺ 17], [RH 17]

... Project Management

- **Follow-up projects**
 - UNIBAS
 - *Vitrivr*: content-based multimodal search
 - *City-Stories*: search in multimedia collections of cultural heritage content, temporal multimedia browser (funded by the Hasler Foundation, CH)
 - *SportSense*: retrieval in sports, in particular soccer, videos
 - UMONS
 - *DigiMIR (2016-2020)* – EU ERDF: Multimedia Information Retrieval for the Creative Industries
 - *DeepSport (2017-2021)* – Walloon Region: Machine Learning for Automating the Analysis and Production of Sports Videos.
 - Koç
 - *SoccerSearch*: multimodal soccer match retrieval

IMOTION Publications (1/7)

- [RGG⁺ 18] Luca Rossetto, Ivan Giangreco, Ralph Gasser, Heiko Schuldt: *Competitive Video Retrieval with vitrivr*. In: Proceedings of the International Conference on Multimedia Modeling (MMM), Bangkok, Thailand, February 2018.
- [RS 18] Luca Rossetto, Heiko Schuldt: *The Long Tail of Web Video*. In: Proceedings of the 24th International Conference on Multimedia Modeling, Bangkok, Thailand, February 2018.
- [DAB⁺ 18] Stéphane Dupont, Ozan Can Altiok, Aysegül Bumin, Ceren Dikmen, Ivan Giangreco, Silvan Heller, Emre Külâh, Gueorgui Pironkov, Luca Rossetto, Yusuf Sahillioglu, Heiko Schuldt, Omar Seddati, Yusuf Setinkaya, Metin Sezgin, Claudiu Tanase, Emre Toyân, Sean Wood, Doguhan Yeke: *VideoSketcher: Innovative Query Modes for Searching Videos through Sketches, Motion and Sound*. In: Proceedings of the eNTERFACE 2015 Workshop on Intelligent Interfaces, Mons, Belgium, January 2018.
- [RGT⁺ 17b] Luca Rossetto, Ivan Giangreco, Claudiu Tănase, Heiko Schuldt: *Multimodal Video Retrieval with the 2017 IMOTION System*. In: International Conference on Multimedia Retrieval (ICMR 17), Bucharest, Romania, June 2017.
- [RS⁺ 17] Luca Rossetto, Heiko Schuldt: *Web Video in Numbers - An Analysis of Web-Video Metadata*. In: arXiv cs.MM, 2017/7
- [RGG⁺ 17] Luca Rossetto, Ivan Giangreco, Ralph Gasser, Heiko Schuldt: *Content-based Multimedia Retrieval using vitrivr*. In: ACM SIGMM Records, 2017/12.

IMOTION Publications (2/7)

- [GRTS 17] Ivan Giangreco, Luca Rossetto, Claudiu Tănase, Heiko Schuldt: *vitivr: A Multimedia Search System supporting Multimodal Interactions*. In: International Conference on Multimodal Communication (ICMC 17), Osnabrück, Germany, June 2017.
- [GGR⁺ 17] Prateek Goel, Ivan Giangreco, Luca Rossetto, Claudiu Tănase, Heiko Schuldt: *Hey, vitivr! - A Multimodal UI for Video Retrieval*. In: Proceedings of the 39th European Conference on Information Retrieval (ECIR 2017), Aberdeen, Scotland, UK, April 2017.
- [RGT⁺ 17] Luca Rossetto, Ivan Giangreco, Claudiu Tănase, Heiko Schuldt, Stéphane Dupont, Omar Seddati: *Enhanced Retrieval and Browsing in the IMOTION System*. In: Proceedings of the 23rd International Conference on Multimedia Modeling, Reykjavik, Iceland, January 2017. **1st place in the competition.**
- [RGT⁺ 16b] Luca Rossetto, Ivan Giangreco, Claudiu Tănase, Heiko Schuldt: *vitivr - A Flexible Retrieval Stack Supporting Multiple Query Modes for Searching in Multimedia Collections*. In: Proc. of the 2016 ACM on Multimedia Conference, Amsterdam, NL, October 2016.
- [TRG⁺ 16a] Claudiu Tănase, Luca Rossetto, Ivan Giangreco, Heiko Schuldt, Stéphane Dupont and Omar Seddati. *The IMOTION System at TRECVID 2016: The Ad-Hoc Video Search Task*. In: Proceedings of the 2016 TRECVID Ad-Hoc Video Search Task, October 2016.

IMOTION Publications (3/7)

- [CSB⁺ 16] Claudiu Cobârzan, Klaus Schoeffmann, Werner Bailer, Wolfgang Hürst, Adam Blažek, Jakub Lokoč, Stefanos Vrochidis, Kai Uwe Barthel, Luca Rossetto: *Interactive video search tools: a detailed analysis of the video browser showdown 2015*. Multimedia Tools and Applications, July 2016.
- [YSC⁺ 16] Kemal Tugrul Yesilbek, Cansu Sen, Serike Cakmak, Tevfik Metin Sezgin: *SVM-Based Sketch Recognition: Which Hyperparameter Interval to Try*, In: Proceedings of the workshop on Sketch-Based Interfaces and Modeling (SBIM '15), pages 117-121, June 2015.
<http://dl.acm.org/citation.cfm?id=2810218>
- [SDM 16a] Omar Seddati, Stéphane Dupont, Saïd Mahmoudi: *Réseaux de neurones convolutionnels profonds pour la reconnaissance d'action dans les vidéos*, CORESA 2016 (COmpression et REprésentation des Signaux Audiovisuels), Nancy, FR, 19-20 May 2016.
- [SDM 16b] Omar Seddati, Stéphane Dupont, Saïd Mahmoudi. *DeepSketch 2: Deep convolutional neural networks for partial sketch recognition*. In: Content-Based Multimedia Indexing (CBMI), 2016 14th International Workshop on. IEEE, 2016. p. 1-6.
- [SDM 16c] Omar Seddati, Stéphane Dupont, Saïd Mahmoudi. *DeepSketch2Image: Deep Convolutional Neural Networks for Partial Sketch Recognition and Image Retrieval*. In: Proceedings of the 2016 ACM on Multimedia Conference. ACM, 2016. p. 739-741.

IMOTION Publications (4/7)

- [TGR⁺ 16] Claudiu Tănase, Ivan Giangreco, Luca Rossetto, Heiko Schuldt, Omar Seddati, Stéphane Dupont, Ozan Can Altıok, Metin Sezgin: *Semantic Sketch-Based Video Retrieval with Autocompletion*. In: Proceedings of the 21st ACM International Conference on Intelligent User Interfaces (IUI'16), Sonoma, CA, USA, March 2016.
- [GS 16] Ivan Giangreco, Heiko Schuldt: *ADAM_{pro}: Database Support for Big Multimedia Retrieval*. Datenbank-Spektrum, Special Issue on Big Data & Information Retrieval, Springer, March 2016.
- [RGH⁺ 16] Luca Rossetto, Ivan Giangreco, Silvan Heller, Claudiu Tănase, Heiko Schuldt, Stéphane Dupont, Omar Seddati, Metin Sezgin, Ozan Can Altıok, Yusuf Sahillioğlu: *IMOTION – Searching for Video Sequences using Multi-Shot Sketch Queries*. In: Proceedings of the 22nd International Conference on Multimedia Modeling (MMM 2016) - Video Browser Showdown Track, Miami, FL, USA, January 2016. Springer LNCS, Vol. 9517, pp. 377-382.
- [RGT⁺ 16] Luca Rossetto, Ivan Giangreco, Claudiu Tănase, Heiko Schuldt, Stéphane Dupont, Omar Seddati, Metin Sezgin, Yusuf Sahillioğlu: *iAutoMotion – an Autonomous Content-based Video Retrieval Engine*. In: Proceedings of the 22nd International Conference on Multimedia Modeling (MMM 2016) - Video Browser Showdown Track, Miami, FL, USA, January 2016. Springer LNCS, Vol. 9517, pp. 383-387.

IMOTION Publications (5/7)

- [RGH⁺ 16] Luca Rossetto, Ivan Giangreco, Silvan Heller, Claudiu Tănase, Heiko Schuldt: *Searching in Video Collections using Sketches and Sample Images – The Cineast System*. In: Proceedings of the 22nd International Conference on Multimedia Modeling (MMM 2016), Miami, FL, USA, January 2016. Springer LNCS, Vol. 9517, pp. 336-341. **Best Demo Award**.
- [RTS 16] Luca Rossetto, Claudiu Tănase, Heiko Schuldt: *Dealing with ambiguous Queries in Multimodal Video Retrieval*. In: Proceedings of the 22nd Int'l Conf. on Multimedia Modeling (MMM 2016), Miami, FL, USA, January 2016. Springer LNCS, Vol. 9517, pp. 898-909.
- [YSC⁺ 15] Kemal Tugrul Yesilbek, Cansu Sen, Serike Cakmak, Tevfik Metin Sezgin: *SVM-Based Sketch Recognition: Which Hyperparameter Interval to Try*, In: Proceedings of the workshop on Sketch-Based Interfaces and Modeling (SBIM '15), pages 117-121, June 2015.
- [AYS16] Ozan Can Altıok, Kemal Tugrul Yesilbek, Tevfik Metin Sezgin: *What Auto-Completion Tells Us About Sketch Recognition*. In: Proceedings of the Sketch-Based Interfaces and Modeling Symposium, Lisbon, May 2016.
- [SKP⁺ 15] Omar Seddati, Emre Kulah, Gueorgui Pironkov, Stéphane Dupont, Saïd Mahmoudi, Thierry Dutoit: *UMONS at MediaEval 2015 Affective Impact of Movies Task including Violent Scenes Detection*, In: Working Notes Proceedings of the MediaEval 2015 Workshop, Wurzen, Germany, 14-15 September 2015.

IMOTION Publications (6/7)

- [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 21st MultiMedia Modelling Conference (MMM2015) - Video Search Showcase Track, Sydney, Australia, pp. 255-260, January 2015, Springer LNCS, Vol. 8936.
- [SDM 15] Omar Seddati, Stéphane Dupont, Said Mahmoudi: *Deepsketch: deep convolutional neural networks for sketch recognition and similarity search*, in Proceedings of 13th International Workshop on Content-Based Multimedia Indexing (CBMI), Prague, Czech Republic, 10-12 June 2015.
- [GAS 14a] Ivan Giangreco, Ihab Al Kabary, Heiko Schuldt: *ADAM — A Database and Information Retrieval System for Big Multimedia Collections*. In: Proceedings of the 3rd 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 37th 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.

IMOTION Publications (7/7)

- [SGS 14] Fabio Sulser , Ivan Giangreco, Heiko Schuldt: *Crowd-based Semantic Event Detection and Video Annotation for Sports Videos*. In: Proceedings of the 3rd International ACM Workshop on Crowdsourcing for Multimedia, Orlando, USA, pp. 63-68, November 2014.
- [SDM 16a] Omar Seddati, Stéphane Dupont, Saïd Mahmoudi: *Réseaux de neurones convolutionnels profonds pour la reconnaissance d'action dans les vidéos*. In: Proceedings of 18^{ème} édition CORESA 2016 (COmpression et REprésentation des Signaux Audiovisuels), Nancy, France, May 2016.
- [SDM 16b] Omar Seddati, Stéphane Dupont, Saïd Mahmoudi: *DeepSketch 2: Deep Convolutional Neural Networks for Partial Sketch Recognition*. In: Proceedings of the 14th International Workshop on Content-Based Multimedia Indexing (CBMI), Bucharest, Romania, June 2016.
- [DAB⁺ 15] Stéphane Dupont, Ozan Can Altıok, Aysegül Bumin, Ceren Dikmen, Ivan Giangreco, Silvan Heller, Emre Külâh, Gueorgui Pironkov, Luca Rossetto, Yusuf Sahillioğlu, Heiko Schuldt, Omar Seddati, Yusuf Setinkaya, Metin Sezgin, Claudiu Tănase, Emre Toyran, Sean Wood, Doguhan Yeke: *VideoSketcher: Innovative Query Modes for Searching Videos through Sketches, Motion and Sound*. In: Proceeding the 11th International one-month Summer Workshop on Multimodal Interfaces (eNTERFACE'2015), Mons, Belgium (in press).

Thank you!

Questions?