# Efficient video retrieval using complex sketches and exploration based on semantic descriptors

**Adam Blažek**
Jakub Lokoč (supervisor)

**Charles University in Prague**
Faculty of Mathematics and Physics

## Introduction

With advance of video recording devices, large video collections are becoming common, yet their sizes, contents and applications are indeed diverse. Hence, the need of effective video search, organization, management and browsing tools arises.

In this work, we propose, implement and evaluate a tool for video search and browsing – Sketch-based Video Browser.

## Feature Signatures

The key-stone of our approach is the feature signatures descriptor. Feature signatures flexibly capture the color distribution of video key-frames and at the same time are comprehensible for users (Fig. 1). Feature signatures are extracted with an adaptive variant of k-means clustering of key-frame pixels. Users are enabled to draw simple color-position sketches (Fig. 2) which might are matched to the extracted feature signatures.

**Figure 1**: A key-frame and the visualization of its feature signature.

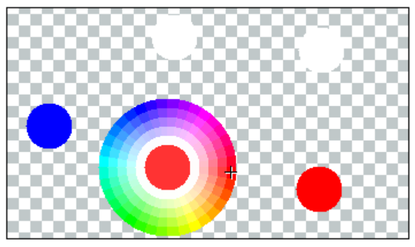

**Figure 2**: Color-Position sketch.

## Results

We carried out a number of user-studies steadily confirming viability of our approach as even novice users were able to search the video database with success. Furthermore, we won an interactive competition in video search Video Browser Showdown[12] in 2014 and 2015. Finally, we published several publications at international conferences and journals regarding this research[1-7].

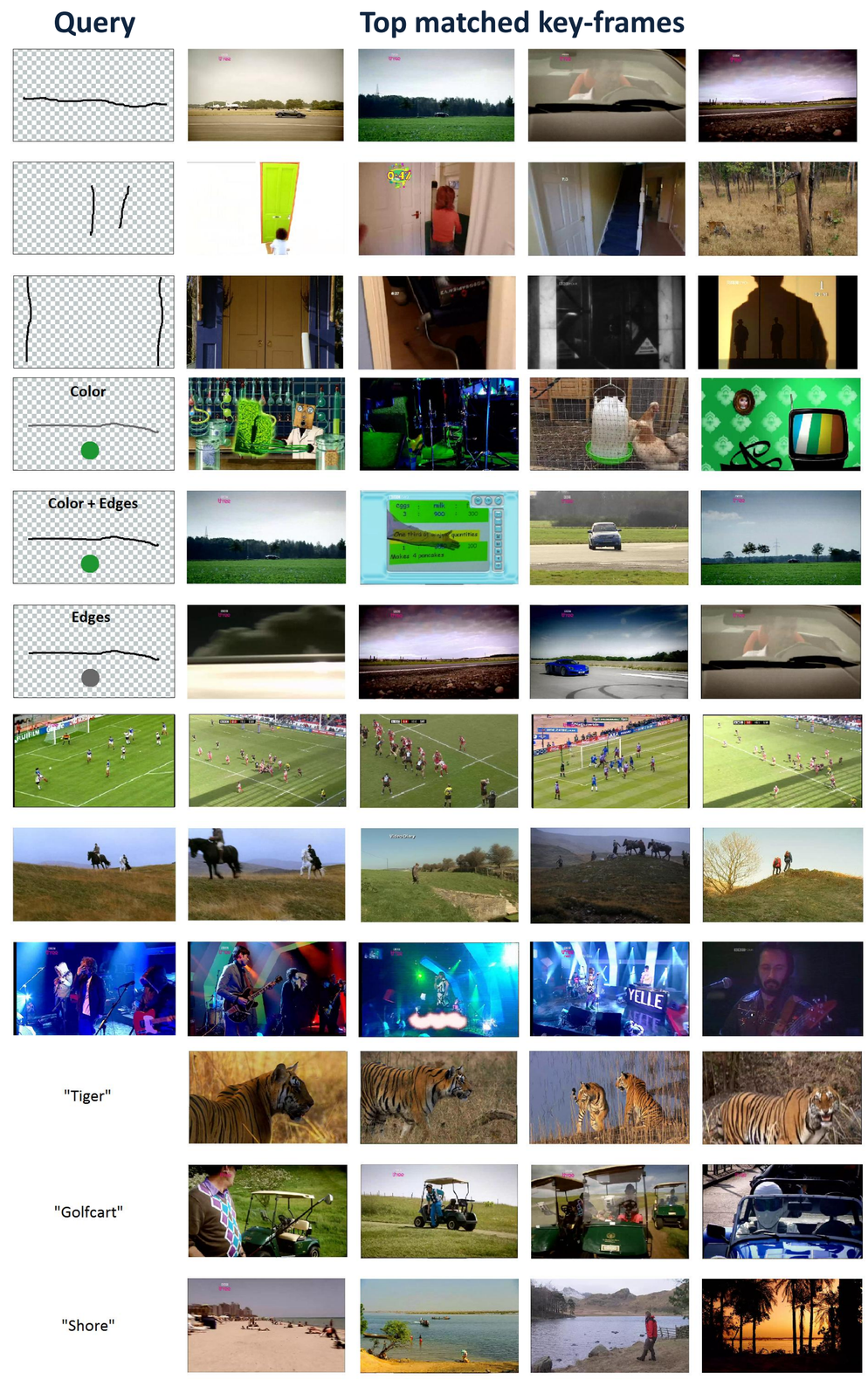


**Figure 3**: Example queries (left) and the respective best matching key-frames (rest of the rows).


## Multimodal Sketches

Additionally, we have incorporated also edge-sketch based video retrieval. During the preprocessing, we compute both local and global edge histograms[8]. Users may retrieve approximate k nearest neighbors from defined edge sketches.

Color and edge sketches may be combined in our multi-modal sketching canvas. Users may put an emphasis on particular modality by setting its importance.

## Semantic Descriptors

Aside from low-level descriptors, it is desirable to capture the actual semantics of the key-frames. We achieve that with already pre-trained deep neural networks[9] which are able to classify the key-frames to no less than 1000 different concepts. Subsequently, the concepts are generalized with WordNet[10] hierarchy and stored in inverted-file index. This allows convenient video retrieval with textual queries.

Furthermore, deep convolutional activation features (DeCAF) are extracted and utilized in semantic similarity search. Thanks to utilization of M-Index[11] data structure we are able to retain real-time responses even for hundreds of hours of video content.

## Our Publications

**[1]** Jakub Lokoč, Adam Blažek, and Tomáš Skopal. Signature-Based VideoBrowser, MMM '14 **[2]** Jakub Lokoč, Adam Blažek, and Tomáš Skopal. On efective known item video search using feature signatures. ICMR '14, **[3]** Adam Blažek, Jakub Lokoč, and Tomáš Skopal. Video retrieval with feature signature sketches. SISAP **[4]** Adam Blažek, Jakub Lokoč, Filip Matzner, and Tomáš Skopal. Enhanced Signature-Based Video Browser MMM '15 **[5]** David Kuboň, Adam Blažek, Jakub Lokoč, and Tomáš Skopal. Multi-sketch Semantic Video Browser, MMM '16 **[6]** Adam Blažek, David Kuboň and Jakub Lokoč. Known-Item Search in Video Databases with Textual Queries . SISAP '16 **[7]** Claudiu Cobarzan, Klaus Schoemann, Werner Bailer, Wolfgang Hurst, Adam Blažek, Jakub Lokoč et.al.. Interactive video search tools: a detailed analysis of the videobrowser showdown 2015. Multimedia Tools and Applications

## References

**[8]** Dong Kwon Park, Yoon Seok Jeon, and Chee Sun Won. Ecient use of local edge histogram descriptorACM 2000 **[9]** Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, CoRR 2014 **[10]** Christiane Fellbaum. WordNet: An Electronic Lexical Database. Bradford Books, 1998 **[11]** David Novak, Michal Batko, and Pavel Zezula. Large-scale similarity data management with distributed metric indexIPM 2012 **[12]** A user-centric media retrieval competition: The video browser showdown 2012-2014. IEEE MultiMedia 2014