

# Video Hyperlinking

Vasileios Mezaris  
CERTH-ITI  
6th Km. Charilaou-Thermi Road, P.O. Box 60361  
Thermi 57001, Greece  
bmezaris@iti.gr

Benoit Huet  
Eurecom  
Campus SophiaTech, 450 Route des Chappes,  
06410 Biot Sophia Antipolis, France  
benoit.huet@eurecom.fr

## ABSTRACT

This is the tutorial abstract for the “Video Hyperlinking” tutorial, presented as part of the 2014 ACM Multimedia Conference.

## Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

## General Terms

Algorithms, Theory, Experimentation

## Keywords

Video analysis, video hyperlinks, non-linear video, hyper-video, linked television

## 1. INTRODUCTION

The tremendous and continuously accelerated growth in the amount of images and videos on the cloud, together with the widespread availability of a wide range of non-PC end-user connected devices (ranging from smartphones and tablets to internet-enabled TV sets), is changing the ways in which people consume visual content. In this environment, the traditional PC-era visual- or text-based search paradigms do not fully meet the expectations of the users, who find it increasingly difficult to navigate in a sea of disassociated video content and discover scattered content items of various origin that may be able to serve their specific information needs, e.g. more information on a story that was briefly presented earlier in the news, or further footage of an object of desire that they just saw in a video. Inspired by the ubiquitous text-based hyperlinks and the way that text hyperlinking has transformed how people navigate through textual and other information, its analogous for video content, i.e. video hyperlinking, is emerging as a promising approach to making video more easily accessible and consumable [1][2]. This tutorial provides a thorough overview of the state of the art in video hyperlinking approaches.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

*MM'14*, Nov. 3–7, 2014, Orlando, FL, USA.

ACM 978-1-4503-3063-3/14/11.

<http://dx.doi.org/10.1145/2647868.2654853>.

## 2. VIDEO HYPERLINKING AND CHALLENGES

Video hyperlinking is the introduction of links that originate from meaningful fragments of video content (e.g. a video shot) and point to other relevant content (which may be visual content, e.g. another video or a segment of it, or any other form of content, e.g. an audio recording or a relevant Wikipedia article) - just like traditional hyperlinks originate from meaningful parts of a text (e.g. a key-phrase or a name) and point to other related resources. However, the manual insertion of such links in newly-created videos is a form of content curation that very few content creators would be willing to perform and maintain over time, because of being an undeniably tedious process. What is needed for making hypervideo feasible is the development of methods for the automatic or semi-automatic identification of related content and the generation of the corresponding links, transforming the present day's disassociated videos in the cloud into a connected and easy to navigate hypervideo collection. This brings new challenges in automatically processing the visual content and understanding the information it conveys at different granularities, in processing associated audio and textual information, and in intelligently exploiting all these analysis results for creating meaningful video hyperlinks [4]. It also raises important questions concerning the granularities that are most appropriate for decomposing and linking the video content [3].

## 3. TUTORIAL CONTENT AND STRUCTURE

This tutorial introduces the vision of hypervideo, and then focuses on two equally important directions: the video (and associated information) analysis that is needed for supporting video hyperlinking; and, the ways for making use of these imperfect analysis results so as to effectively discover and establish meaningful video links at suitable granularity levels.

The presentation of the material is structured in three main parts.

The introductory part (Part A), Video Hyperlinking Motivation and Vision, briefly introduces what video hyperlinking is, highlights the functionalities we want to achieve with video hyperlinking, and explains why this is a both important and challenging task.

The motivation and vision discussion is followed by the first of the two more technical parts of the tutorial (Part B), Video fragment creation and annotation for hyperlinking. This starts by establishing an analogy between text hyper-

linking and video hyperlinking, in order to explain which video analysis technologies we need for realizing the vision of automatic video hyperlinking and what exact purpose each of them serves. It then goes on to present the state of the art and discuss in more depth selected video analysis techniques. These include techniques for

- Video temporal segmentation to shots
- Video temporal segmentation to scenes
- Visual concept detection
- Event detection

Other potentially relevant technologies, such as duplicate or near-duplicate image detection or object re-detection in video will also be mentioned and briefly commented. The presentation of each of the analysis techniques is not limited to a theoretical study, but includes the presentation of indicative experimental results and demos that clearly illustrate what results these techniques can deliver and how these results contribute to the objective of video hyperlinking. Additional reading for the presented technologies is also listed at the end of each section of the presentation.

In the third part of the tutorial (Part C), having outlined the state of the art in individual video analysis tasks, we discuss how to effectively use and combine their results towards discovering and establishing meaningful video links at suitable granularity levels. Indeed, the granularity at which the various multimodal information is available depends on the modality itself; video frames are displayed 24, 25 or 30 per second while subtitle (or transcript) items, OCR text and other metadata span over many frames, and in some cases (i.e. video title) over the entire video. As part of the process linking related video content, we provide insight on various approaches using text, audio and visual content analysis results, highlighting what does and what doesn't work. The outcome of related video hyperlinking experiments on real data [5] are shown, evaluated on a public dataset and discussed.

We also identify future directions for better addressing the video hyperlinking problem, and conclude by briefly sketching a few hypervideo usage scenarios and applications.

## 4. ACKNOWLEDGMENTS

This work was supported by the European Commission under contract FP7-287911 LinkedTV.

## 5. REFERENCES

- [1] Linkedtv: Television linked to the web.  
<http://www.linkedtv.eu/>.
- [2] Mediaeval benchmarking initiative for multimedia evaluation: Search and hyperlinking task.  
<http://www.multimediaeval.org/mediaeval2014/hyper2014/>.
- [3] E. Apostolidis, V. Mezaris, M. Sahuguet, B. Huet, B. Cervenkova, D. Stein, S. Eickeler, J.-L. R. Garcia, R. Troncy, and L. Pikora. Automatic fine-grained hyperlinking of videos within a closed collection using scene segmentation. In *Proc. ACM Multimedia 2014*, Orlando, FL, USA, November 2014.
- [4] L. Nixon, V. Mezaris, and J. Thomsen. Seamlessly interlinking tv and web content to enable linked television. In *ACM Int. Conf. on Interactive*

*Experiences for Television and Online Video (TVX 2014)*, Adjunct Proceedings, Newcastle Upon Tyne, UK, June 2014.

- [5] M. Sahuguet, B. Huet, B. Cervenkova, E. Apostolidis, V. Mezaris, D. Stein, S. Eickeler, J. L. Redondo Garcia, R. Troncy, and L. Pikora. LinkedTV at MediaEval 2013 search and hyperlinking task. In *MEDIAEVAL 2013, Multimedia Benchmark Workshop, October 18-19, 2013, Barcelona, Spain*, Barcelona, SPAIN, 10 2013.