

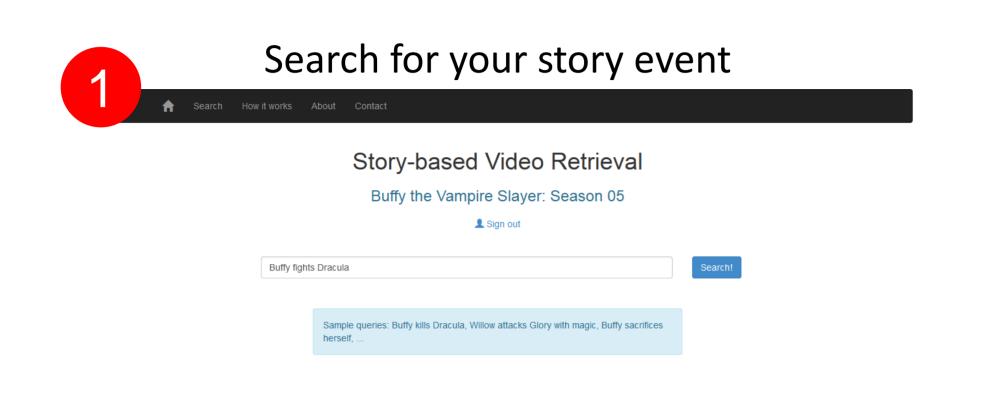
Story-based Video Retrieval in TV series using Plot Synopses

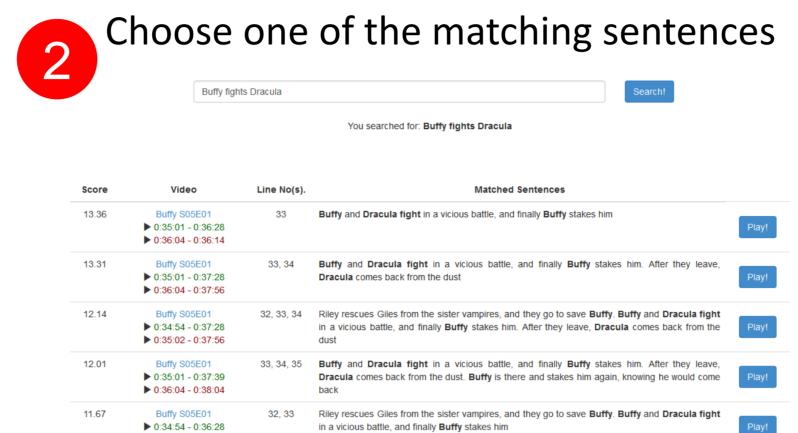
Makarand Tapaswi, Martin Bäuml, Rainer Stiefelhagen

Computer Vision for Human Computer Interaction, Karlsruhe Institute of Technology, Germany

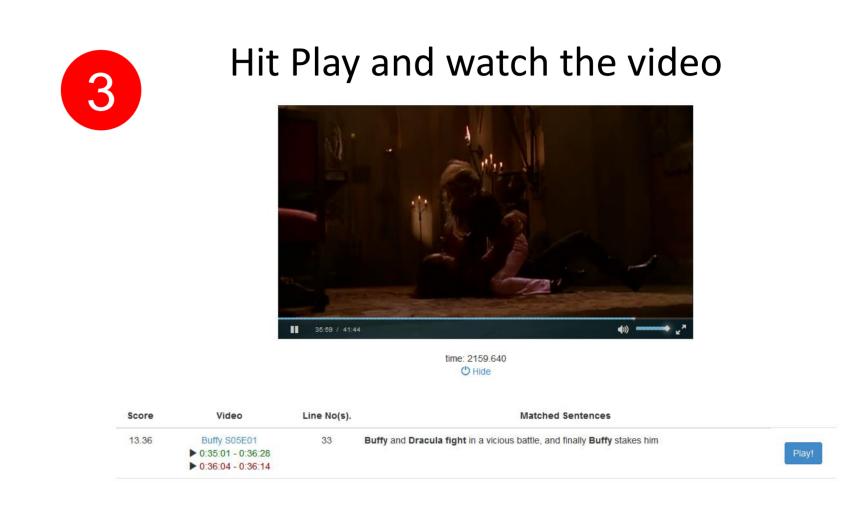
Contributions

- Automatic retrieval of video snippets related to events in a storyline
- Alignment of human written natural language descriptions (plot synopses) to shots of a video
- Annotations for alignment of text to video for complete season 5 of Buffy the Vampire Slayer





0:35:02 - 0:36:14



Giles has Willow start scanning books into a computer so that they can be resources for the

Buffy seems distracted and after hearing about the truths of Dracula, she leaves abruptly

Motivation

While concept-based video retrieval is a big step forward from low-level content-based search, it searches only for mid-level events. In this work, we propose to use crowd-sourced information such as plot synopses on Wikipedia for structured videos (TV series, movies) to augment the video information and allow story-related queries within a long storyline.

Text Processing

Part-of-speech tagging

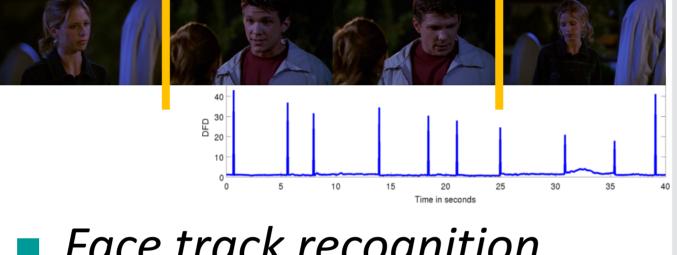
Buffy/NNP awakens/VBZ to/TO find/VBP Dracula/NNP in/IN her/PRP bedroom/NN ./. She/PRP is/VBZ helpless/JJ against/IN his/PRP powers/NNS and/CC unable/JJ to/TO stop/VB him/PRP from/IN biting/VBG her/PRP ./.

Coreference resolution

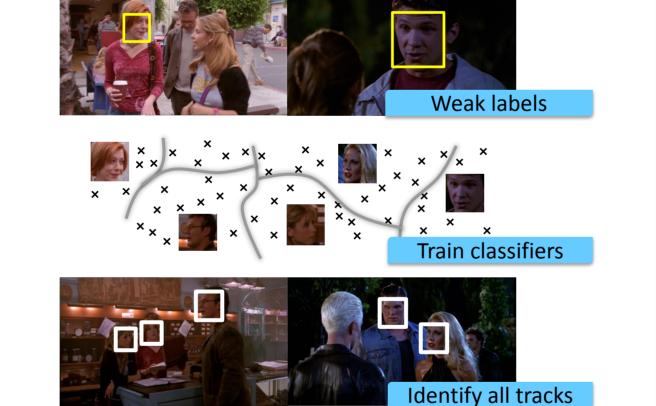
Buffy awakens to find **Dracula** in her bedroom. She is helpless against his powers and unable to stop him from biting her. When she wakes the next morning ...

Video Processing

Shot detection



Face track recognition



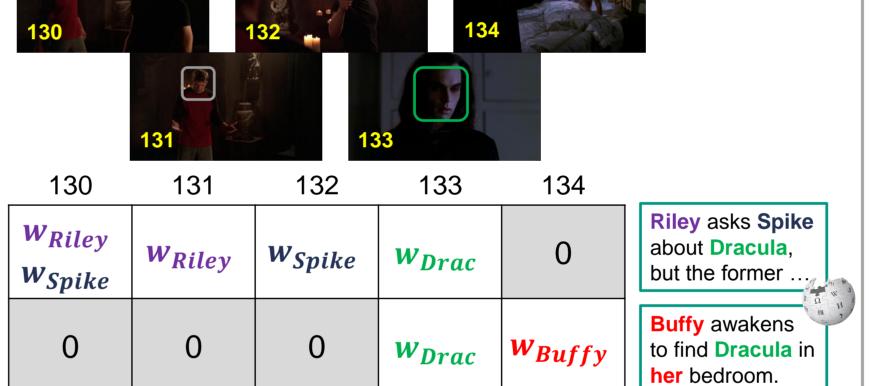
Text-Video Alignment

Idea

Compute a similarity score between every shot and every sentence.

Cues



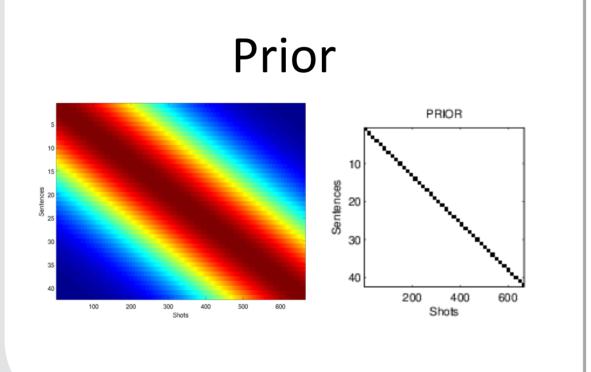


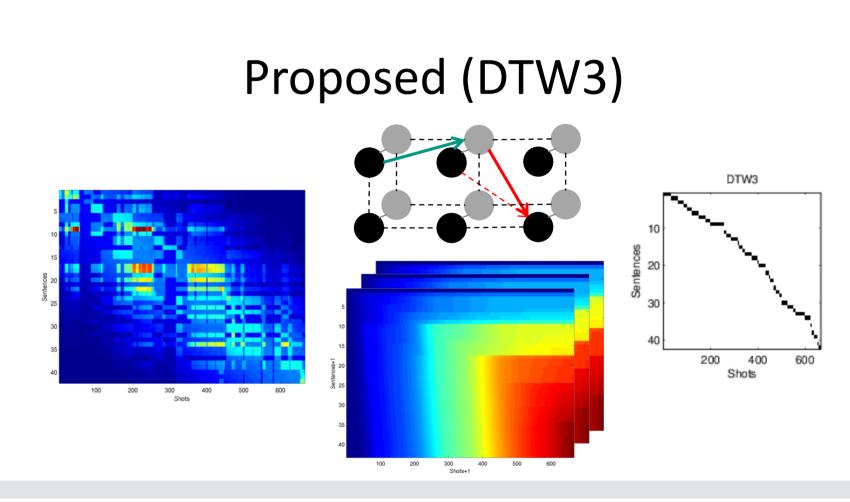


Shots

Obstinate bloody machine start **scanning** simply refused to work. books into a computer so there 00:04:30,093 --> 00:04:32,527 Let's get **scanning**. can be resources I wanna see this puppy go. for the gang to use. 00:05:38,493 --> 00:05:43,521 We're doing this because I He then tells her want you to have everything you need at that he's **going** to fingertips. **England** because it 00:05:43,613 --> 00:05:45,843 seems he's no You see, I'm going back to longer needed by England. Buffy. 00:05:47,973 --> 00:05:50,123 You're what?

Alignment methods



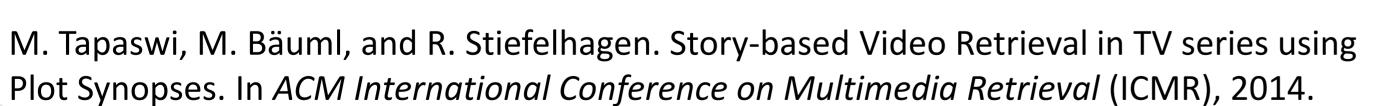


Data and Story-events

- Buffy the Vampire Slayer (season 5, 22 episodes)
- Data statistics
 - 15700+ shots
 - 800+ sentences
 - 21000+ face tracks
 - Per episode: 720 shots, 36 sentences
- Story events
 - 62 queries (fan-forum + annotators)







Contact

{tapaswi, baeuml}@kit.edu

Project page (data and code)

http://cvhci.anthropomatik.kit.edu/projects/mma





Results

Dracula talks to **Buffy** of all the things **he** will do for **her** while **she** struggles to regain control of **herself**. Xander tries to stop Riley from going after Dracula, but Riley knocks him out with one punch.

BF-1:29. Giles finds himself victim to the Three Sisters who effectively keep him distracted.





Alignment Evaluation (fraction of shots assigned to correct sentence)

Method		Buffy E01	Buffy E02	Buffy E03	Buffy E04	Average E01 - E22
Human		81.5	86.4	77.5	72.8	_
Prior		2.9	23.8	27.9	8.8	10.11
Character ID	DTW3	42.2	43.8	40.4	40.3	41.17
Subtitles	DTW3	20.4	48.4	35.3	30.1	37.00
Char-ID+Subt.	DTW3	40.8	51.3	41.4	47.6	49.16

Acknowledgment: This work was supported by the Quaero Programme, funded by OSEO.



