Annotated Dataset for Sub-Shot Segmentation Evaluation

Description

This dataset was created for evaluating the performance of the developed motion-driven approach for fine-grained temporal segmentation of user-generated videos, that is reported in [1].

Given the fact that user-generated videos are most commonly captured without interruption using a single camera, thus being single-shot videos, their fine-grained temporal segmentation aims to identify visually coherent parts (called sub-shots) that correspond to individual actions taking place during the video recording, such as left camera panning, camera zoom-out or tracking of a moving object.

Based on the above, the created dataset contains the ground-truth sub-shot segmentation for 33 single-shot videos. This ground-truth segmentation was created by human annotation of the sub-shot boundaries for each video, where each boundary indicates the end of a visually and temporally contiguous activity of the video recording device and the start of the next one (e.g. a downward camera tilting that is followed by a camera zoom-in). Overall, our dataset contains 674 sub-shot transitions.

The set of videos can be divided in three parts:

1. *Own Videos*: 15 single-shot videos of total duration 6 minutes, which contain clearly defined fragments that correspond to several video recording activities;
2. *Amateur Videos*: 5 single-shot amateur videos of total duration 17 minutes, found on the YouTube platform;
3. *Movie Excerpts*: 13 single-shot parts of known movies of total duration 46 minutes which represent professional video content.

The videos of the first part were recorded in the external spaces of our (CERTH-ITI) facilities using an iPhone 5 smart-phone. These videos and their ground-truth are freely provided through the "Provided files" section below.

The videos of the second and third part can be found on the YouTube platform (links in the "Video collection" section below). For these videos we only provide the ground-truth data again through the "Provided files" section below.

Video collection

In the following table, for each video we provide:

- The video ID;
- The video filename for "Own videos" and the YouTube video name and URL for "Amateur videos" and "Movie Excerpts" (accessed on 12/11/2017);
- The video frame-rate that directly affects the ground-truth annotation; the videos of the second and third part should be downloaded at the corresponding frame-rates (default values when using the KeepVid web application - https://keepvid.com/) to match the created annotations.

ID	Provided video filename or YouTube link	Frame rate
Own Videos		
own_1	vid_1.mp4	29.97
own_2	vid_2.mp4	29.97
own_3	vid_3.mp4	29.97
own_4	vid_4.mp4	29.97
own_5	vid_5.mp4	29.97
own_6	vid_6.mp4	29.97
own_7	vid_7.mp4	29.97
own_8	vid_8.mp4	29.97
own_9	vid_9.mp4	29.97
own_10	vid_10.mp4	29.97
own_11	vid_11.mp4	29.97
own_12	vid_12.mov	29.97
own_13	vid_13.mp4	29.97
own_14	vid_14.mp4	29.97
own_15	vid_15.mp4	29.97
Amateur Videos		
amateur_continuous_shot_music_video	"Continuous Shot Music Video" (https://youtu.be/N0o01jnsUQg)	29.97
amateur_fearless_5_year_old	"Fearless 5 year old uses bow and arrow to pull out own tooth!" (https://youtu.be/u4yoKwARLwI)	24.87
amateur_hendrickson_high_school	"Hendrickson High School Lip Dub 2014" (https://youtu.be/z6B2L-Paw8U)	29.97
amateur_kid_pulls_out_own_tooth	"Kid pulls out own tooth with bow and arrow (happy)" (https://youtu.be/oD5CBgtErOk)	30
amateur_shake_it_off	"Shake It Off Lip Dub - Delta Sigma Phi" (https://youtu.be/s2YHHIOocxs)	29.97
Movie Excerpts		
movie_excerpt_atonement	"Atonement - Dunkirk Scene, Five minute single take" (https://youtu.be/QijbOCvunfU)	30
movie_excerpt_boogie_nights_1	"Boogie Nights 1-Long Take" (https://youtu.be/8WEONgYsk1Q)	24
movie_excerpt_boogie_nights_2	"Boogie Nights - Do Your Thing" (https://youtu.be/PybwHMIxIDk)	29.97
movie_excerpt_children_of_men_1	"Children of Men 2006 Long Take 3" (https://youtu.be/MjFHqohaHYU)	23.97
movie_excerpt_children_of_men_2	"Children of men: Car Scene" (https://youtu.be/QfBSncUspBk)	29.97
movie_excerpt_goodfellas	"Goodfellas - Steadicam Shot" (https://youtu.be/3Sr-vxVaY_M)	30
movie_excerpt_hanna	"Hanna long shot" (https://youtu.be/KXIpgg58xDg)	23.97
movie_excerpt_magnolia	"135 second shot from Magnolia" (https://youtu.be/dDMdzg_QmxM)	25

ID	Provided video filename or YouTube link	Frame rate
movie_excerpt_raging_bull	“RAGING BULL - Title Fight entrance” (https://youtu.be/9b6L1uq_bik)	29.97
movie_excerpt_spectre	“Spectre - Opening Tracking Shot in 1080p” (https://youtu.be/cbqv1kbsNUY)	23.97
movie_excerpt_the_protector	“The Protector 2005 Long take” (https://youtu.be/zESe7U467vs)	23.97
movie_excerpt_the_shinning	“The Shining - steadishot by Garret Brown” (https://youtu.be/cy7ztJ3NUMI)	25
movie_excerpt_touch_of_evil	“The Long Take: Touch of Evil” (https://youtu.be/RaN4R6KRSY0)	29.97

Provided files

The needed files for using the created dataset can be downloaded from here. After unpacking the compressed file ("dataset.zip") a structure of directories will be generated. In this structure:

1. The "videos/own_videos" directory contains: a) the video files of the first part of the dataset, and b) the "list.txt" file which lists the video id, the filename and the frame-rate for each video of the "Own videos" collection, in a tab-separated format.
2. The "videos/amateur_videos" directory contains the "list.txt" file which lists the video id, the filename and the frame-rate for each video of the "Amateur videos" collection, in a tab-separated format.
3. The "videos/movie_excerpts" directory contains the "list.txt" file which lists the video id, the filename and the frame-rate for each video of the "Movie excerpts" collection, in a tab-separated format.
4. The "ground_truth" directory contains 33 txt files (one per video) where each file is named after the ID of the corresponding video and includes the annotation of the sub-shots in the video; each line of this file corresponds to the starting frame of a sub-shot of the video (using a zero-based index of frames).
5. The "evaluation" directory contains a simple Matlab script, called "eval_segm.m", that was used for evaluating the results of the sub-shot segmentation analysis.

Copyright

CERTH-ITI holds the copyright of the “*Own videos*” part of the dataset. These videos can be freely used for academic/research purposes only and their public playback is not allowed. CERTH-ITI does not own the copyright of the “*Amateur Videos*” and “*Movie Excerpts*” parts of the dataset. The use of these videos must abide by the YouTube copyright policy (https://www.youtube.com/yt/about/copyright/). Any users of these videos must accept full responsibility for the use of these parts of the dataset, including but not limited to the use of any copies of copyrighted videos that they may create from the dataset.

Citation

If you use this dataset, please cite the following scientific work:

[1] K. Apostolidis, E. Apostolidis, V. Mezaris, *"A motion-driven approach for fine-grained temporal segmentation of user-generated videos"*, Proc. 24th Int. Conf. on Multimedia Modeling (MMM2018), Bangkok, Thailand, Feb. 2018.

```
@InProceedings{kapost_2018_MMM,
author = {Apostolidis, Konstantinos and Apostolidis, Evlampios and
Mezaris, Vasileios},
title = {A motion-driven approach for fine-grained temporal
segmentation of user-generated videos},
booktitle = {The 24th International Conference on Multimedia
Modeling (MMM)},
month = {Feb},
year = {2018}
}
```

Contact

For any queries, please contact the authors of the aforementioned scientific work at: kapost@iti.gr (K. Apostolidis), apostolid@iti.gr (E. Apostolidis), bmezaris@iti.gr (V. Mezaris).

Acknowledgement

This work was supported by the EU's Horizon 2020 research and innovation programme under grant agreement H2020-732665 EMMA.