

The Livius Documentation

Author: Parnia Bahar

April 21, 2015

Abstract

This Documentation serves as a short description for all required knowledge and previous works regarding the LiVius project in Max-Planck institute of intelligent system, Empirical Inference department. The main aim of the work is to automatically create an appropriate layout for online courses and lectures. In this case, a proper framework should be composed of both the lecture slides and the speaker/teacher. To this aim, computer vision and machine learning based algorithms are applied.

1 Introduction

There are many advantages to online and computer-based learning when compared to traditional face-to-face courses and lectures, however, there are a few disadvantages as well. Online courses are great for self-motivated individuals who want to learn new skills and advance their careers and their knowledge without attending in the college or university. Getting education online will also save tons of money. Although some would argue that online education is only an awesome alternative to traditional education because of the savings and convenience, there are actually many other advantages. Only online education fully integrates itself into todays educational technology. It is also more efficient for fast and especially motivated learners and offers skills that lack resources in traditional education despite high demand from employers.

Class work can be scheduled around work and family. Reduces travel time and travel costs for off-campus students. Students may have the option to select learning materials that meets their level of knowledge and interest. Students can study anywhere they have access to a computer and Internet connection. Self-paced learning modules allow students to work at their own pace. Flexibility to join discussions in the bulletin board threaded discussion areas at any hour, or visit with classmates and instructors remotely in chat rooms. Instructors and students both report e-Learning fosters more interaction among students and instructors than in large lecture courses. e-Learning can accommodate different learning styles and facilitate learning through a variety of activities. Develops knowledge of the Internet and computers skills that will help learners throughout their lives and careers. Successfully completing online or computer-based courses builds self-knowledge and self-confidence and encourages students to take responsibility for their learning. Learners can test out of or skim over materials already mastered and concentrate efforts in mastering areas containing new. information and/or skills [?]. Therefore, the demand of taking online classes is increasing and there are many universities and educational organizations which offer free online courses.

1.1 The Framework

One can make see the whole framework of the project in the figure 1. The whole project can be divided into three major tasks. The first one is video processing and it is related to all algorithms and pre-processing steps for slide detection and speaker tracking. Here, the processing tasks can be done either on the video file itself or on the separate frames. The second one can be called video editing and it includes all necessary works for the codecs of the audio and video, the final layout, the background image, the logo and concatenating all together to form the final version of the video. The third task is the streaming the prepared video on the Internet.

1.2 Output video layout

In order to broadcast the final video, the full HD size is used. The final video size to stream is 1920x1320. Figure 2 depicts the final layout and the title and background images.

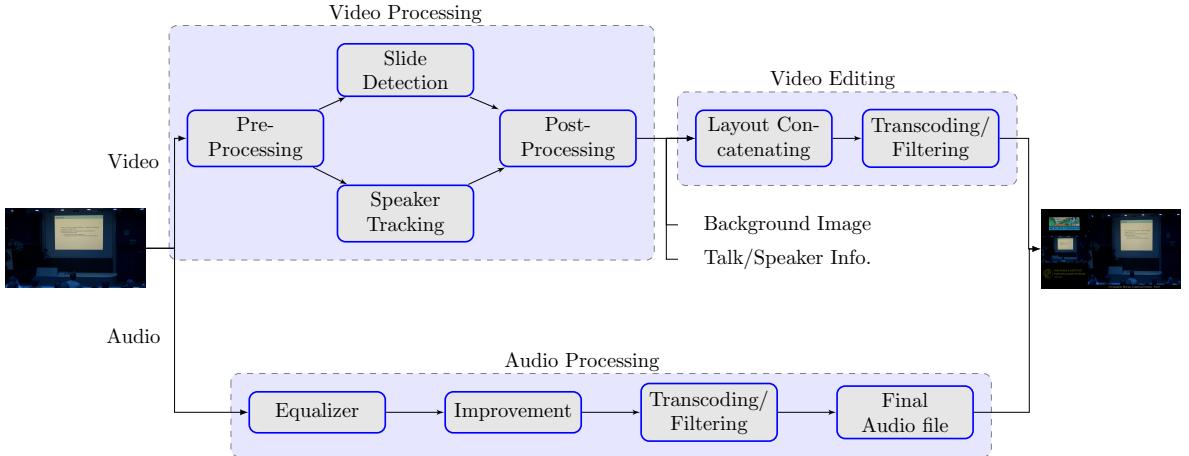


Figure 1: The whole framework



Figure 2: The final video layout

2 Requirements

2.1 Hardware

The current hardware and tools are shown in figure 3. Figure 3c indicates the 4k Blackmagic Camera, its power supply, audio cables and the ultrasonic 24mm lens. Figure 3c is the 2-TB USB3.0 external hard disk for any mobility. Figure 3d shows the 1-TB Solid State Drive (SSD) which suits with Blackmagic camera and 3d includes Sata HDD docking station and reader along with its secondary power supply and USB3.0 port. Besides the shown laptop, there is another desktop PC as an implementation platform with following properties.

- Processor: Intel Core i7-5930K CPU @ 3.50GHz 12
- Memory: 32 GiB
- Graphic: NVIDIA GeForce GTX 750/PCIe/SSE2
- Disk: 3 Data raid for 12 TB
- Operating system: Ubuntu 14.04 and windows 7

2.2 Software

There are two major aspects which should be taken into account in this work. As it was already mentioned the framework needs both video processing and video editing. Since, there is no open source library to handle both tasks, It is better to stay with C++ or Python. To this end, we checked several libraries and tools listed as follow.

1. Nuke production:

A software for composing multiple images and effects, motion graphics, not open source, no good documentations. It seems to be for design, film editing, visualization, etc. <https://www.thefoundry.co.uk/products/nuke/>

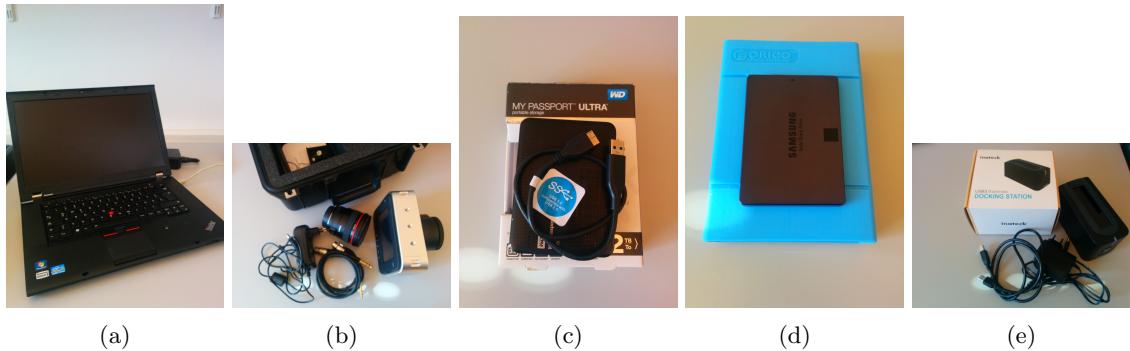


Figure 3: The hardwares

2. MLT framework:

It is an open source tool based on C++ for audio/video editing, media trasncoding, web stream and TV broadcasting. <http://www.mltframework.org/>

3. MoviePy:

MoviePy is a Python module for video editing, which can be used for basic operations (like cuts, concatenations, title insertions), video compositing, or to create advanced effects. It can read and write the most common video formats. <http://zulko.github.io/moviepy/>

In order to develop new algorithms as convenient as possible, it was decided to take advantages from pre-built libraries and use Python language (Python 2.7.6 where /usr/bin/python). The required libraries and packages have been listed as below.

- ffmpeg
- Moviepy
- OpenCV-2.4.11

For more details concerning the list of all available python modules, one can use “help(‘modules’) ” in the Python prompt.

3 Initial Attempts

3.1 Camera Setting

4k Blackmagic camera has different coding formats to store the recorded data and it supports two full HD sizes such as 4k and 2k video . The 4k has 1920x1080 pixels while the size of the 4f video recording is 3840x2160. For each one, there are four different codecs. The codecs are ProRes Proxy, ProRes LT, ProRes 422 and ProRes HQ so that the first one has the lowest quality and the last one takes the highest. Figure 5 shows the main menu and the sub-menus of the 4k blackmagic camera. This camera has two disk formats like HFS+ and exFAT (Figure 4b). The former supports Mac OS and the latter works for windows. There in no format for Linux. The most important part of the menu is setting by which one is able to specify various adjustments. The setting part has 4 different sections.

1. Camera (figure 4c)

Here, it is possible to adjust camera Id, date, time, ISO, white balance and shutter angle. In our records, the ISO, white balance and shutter angle have been adjusted into 400, 5200k and 180° respectively.

2. Audio (figure 4d)

The second section is related to audio adjustments and two microphones inputs.

3. Recording (figure 4e)

The third part supports the recording task and provides different options. In recording format, one can choose either the raw data or one of the 8 codecs. Here, one might determine the fps and time lapse. In our experiments, we use 30 fps.

4. Display

It includes some adjustments corresponding to the screen display of camera like brightness and so on.

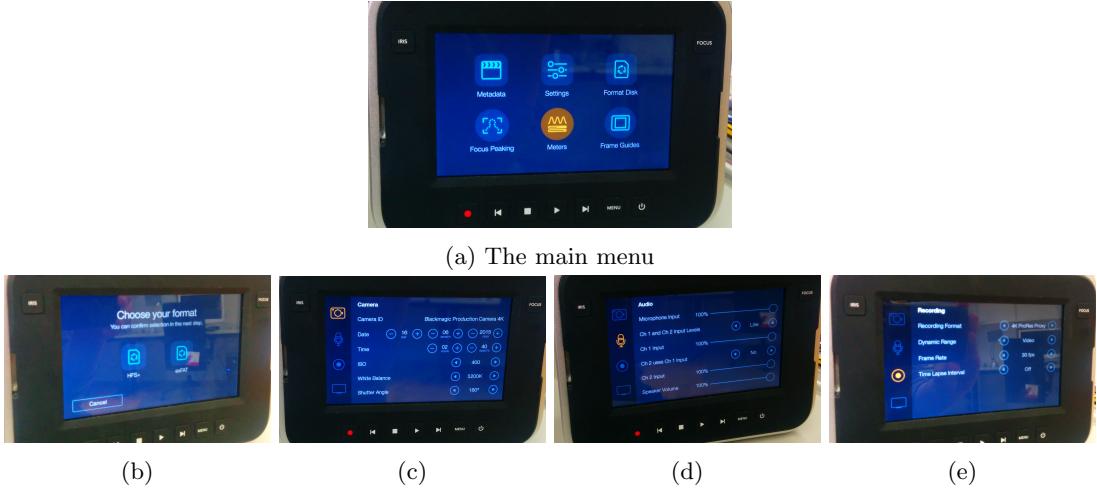


Figure 4: The setting menu of Blackmagic camera

3.2 Camera Codec Formats

The camera has 8 various recording formats as listed in table 1. After trying them, we reached the conclusion that base on the fact that the bit-depth of the video for all formats is 10 bits and according to the size, $PSNR$ values and video quality, the **4k ProRes Proxy** was chosen to record the videos with 30 frame per second (fps). Moreover, different codings with the appropriate container have been examined as shown in table 2. One can compare the quality of coded videos. The **.mp4** container and **x264** coder might be selected as the most promising ones for the final video.

Table 1: The different codecs of 4k Blackmagic camera

Camera codec formats	Size (MB/sec)	Size (GB/90 min)	$PSNR$
4k ProRes Proxy	23.1	125	20 Db
4k ProRes LT	51.6	279	33 Db
4k ProRes 422	74.0	400	22 Db
4k ProRes HQ	108.7	587	99 Db
2k ProRes Proxy	6.3	34	32 Db
2k ProRes LT	13.7	73	37 Db
2k ProRes 422	19.2	103	33 Db
2k ProRes HQ	28.0	151	99 Db

Table 2: The different codecs of 4k Blackmagic camera versus different kind of containers and transcoders. **X** : bad quality , **✓** : good quality , **∞** : medium quality

Camera codec formats	.webm (vpx)	.mp4 (x264)	.mp4 (mpeg4)	.avi (raw video)	.agv (theora)
4k ProRes Proxy	∞	✓	X	X	X
4k ProRes LT	∞	✓	X	X	X
4k ProRes 422	∞	✓	X	X	X
4k ProRes HQ	∞	✓	X	X	X
2k ProRes Proxy	X	∞	X	X	X
2k ProRes LT	X	✓	X	X	X
2k ProRes 422	∞	✓	X	X	X
2k ProRes HQ	∞	✓	X	X	X

4 Acknowledgments

I would like to express my very great appreciation to all of team members, Michael Hirsch, Senya Polikovsky, Edgar Klenske, Michael Schober and Parnia Bahar for their patient, valuable and constructive suggestions during this research work.

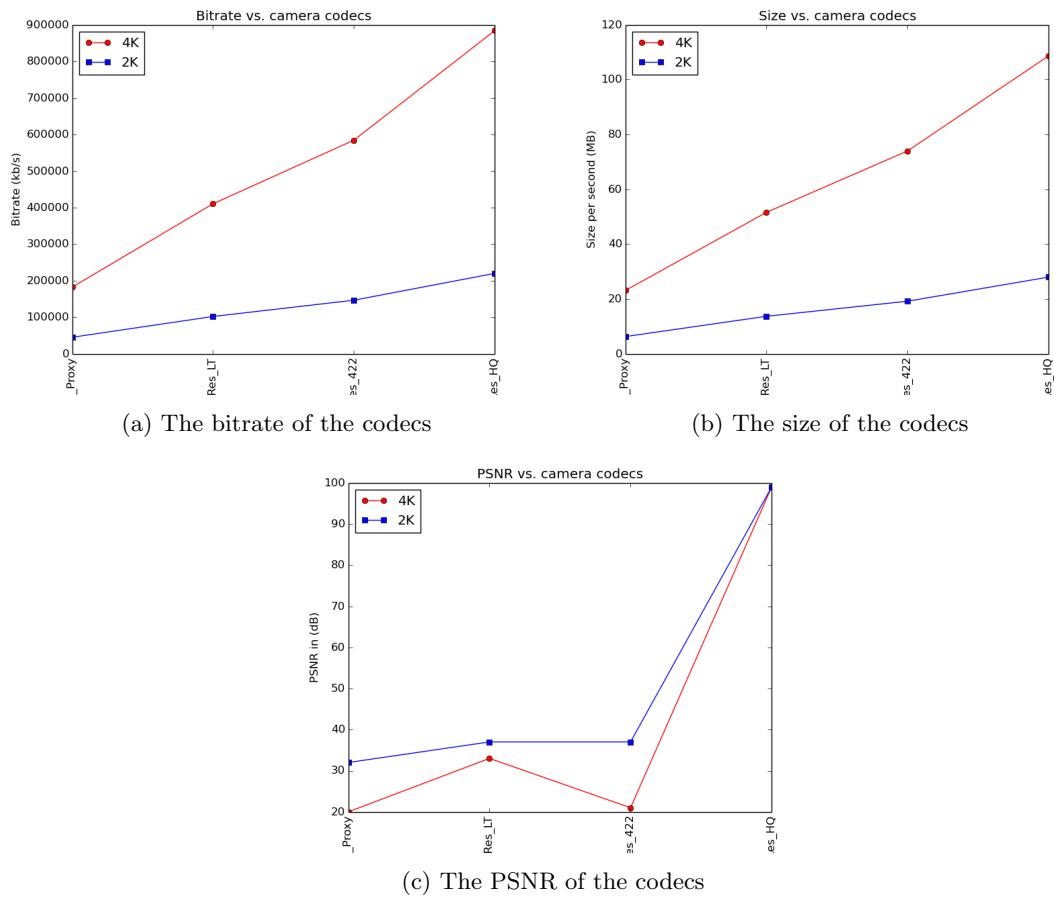


Figure 5: Different comparison of the codecs