Image stabilization for video productions

A survey about technologies and methods for counteracting blurry footage.

Second bachelor thesis

In partial fulfilment of the requirements for the academic degree of
Bachelor of Science in Engineering

in Media Technology
at the University of Applied Sciences St.Pölten

by:
Omar Josef Awad
mt171001

Supervising tutor: Dipl.-Ing. Michael Bock, BSc

Advising tutor abroad: Prof. Michael Harper

[Salzburg, 09.04.2020]

Declaration of Authorship

I hereby declare that,

- the thesis submitted is my own unaided work. All direct or indirect sources used are acknowledged as references.

- I have never submitted this topic within national or foreign territories to an auditor for grading or as an examination paper before.

This thesis is equal to the thesis that was submitted to and graded by the advisor.

Salzburg, 09.04.2020

....

Place, Date Signature

Abstract

Video stabilization is a persistent matter in video productions. With recording systems getting smaller and the amount of video content continually rising, it has become an essential question of how to tackle unwanted shakiness in video footage. Over the years multiple methodologies have been developed and in use that provides a possible solution for these problems. These technologies are continually evolving, and new ones are being introduced. This thesis provides an overview of commonly used technologies, describing how they function and what technologies are currently considered state of the art and how they compare to each other. It also provides an outlook on how the introduction of machine learning has made an impact on this field of research. Furthermore, an experiment with a complementary survey was conducted comparing two in video editing commonly used stabilization tools.

Kurzfassung

Videostabilisierung ist eine anhaltende Thematik für Videoproduktionen. Da Aufnahmesysteme immer kompakter werden und die Menge an Videoinhalten kontinuierlich ansteigt, ist es eine die Branche anhaltend, beschäftigende Frage, wie man ungewünschten Verwacklungen in Videoinhalten beikommen kann. Über die Jahre wurden bereits verschiedenste Technologien entwickelt und angewendet. Diese Technologien entwickeln sich stets weiter und neue werden kontinuierlich vorgestellt. Diese Arbeit gibt einen Überblick darüber, welche Technologien derzeit in der Branche Anwendung finden. Sie beschreibt wie diese Technologien funktionieren und wie sie im Vergleich zu einander stehen. Eine weitere Thematik dieser Arbeit ist auch, wie die Einführung von künstlicher Intelligenz Auswirkungen auf dieses Gebiet hat. Zusätzlich wurde ein Experiment mit zusammenhängender Befragung durchgeführt, in dem zwei in der Branche etablierte digitale Stabilisationsstools verglichen werden.

Inhaltsverzeichnis

Declaration of Authorship II

Abstract III

Kurzfassung IV

Inhaltsverzeichnis V

1 Introduction 1

2 Method 3

2.1 Evaluating existing research 3

2.2 Comparison experiment 4

2.3 Survey about experimental results 6

3 Analysis 7

3.1 Image stabilization during capture 7

3.1.1 Internal image stabilization 7

3.2 External image stabilization technologies 11

3.2.1 Steadicam 12

3.2.2 Electric gimbal systems 13

3.3 Image stabilization in post-production 16

3.3.1 Feature point extraction and image matching 16

3.3.2 Motion estimation and trajectory construction 18

3.3.3 Motion compensation 18

4 Discussion 20

4.1 Stabilization methods used during production 20

4.2 Post – production based stabilization methods 21

4.2.1 Experimental comparison of two post-processing stabilization tools 22

4.2.2 Survey about experimental results 28

5 Conclusion 33

Literature 34

List of Figures 38

List of Tables 39

1 Introduction

This thesis aims to provide an insight into currently existing technologies and technology proposals that cover methods of achieving video stabilization for video productions. Since the amount of and demand for high-quality video content is continually rising, the question of achieving stable, high-quality footage has become an essential aspect of video production workflows. (Zhang et al., 2017) It also mostly has a very negative impact on the result, as most viewers consider unstable footage as unpleasant and said footage might also be unusable for further utilization. (Hsu et al., 2012) Due to involuntary muscle contractions or hand-shakes known as physiological tremors (Smaga, 2003) or other unwanted movements that could occur during capture, the resulting footage often contains an extensive amount of unwanted blur. As cameras are becoming smaller in form factor and able to capture in higher resolutions, such movements lead to even more amplified unwanted results and blurriness. (Sachs et al., 2006) This has led many manufacturers and scientists to develop solutions in order to compensate for these movements. (Golik, 2006) There are some hardware-based solutions that are being applied to counteract shaky footage, but these are often heavy and inflexible for casual capture. In-built stabilization systems are often sufficient for reducing high-frequency jitter. However, they still struggle sometimes with low-frequency disturbances that can occur during walking while recording. This is why there is also a focus on image post-processing solutions since they do not require any hardware and are more sufficiently equipped to handle a broader range of types of shaky motion. (Zhang et al., 2017) Generally speaking, it is possible to separate the stabilization technologies in those which are utilized during the capturing process of the footage and those who are employed in the post-production stage. During the recording stage, most of these solutions are hardware-based trying to counter the movement of the camera body or lenses. However, there are software solutions as well, which try to minimize stutters caused by shaky camera movement digitally. (Liu et al., 2009) All in all, it can be said that dozens of different methods exist to tackle the problem of shaky footage. They differ in the way they work, as well as in their realistic usage scenario. Therefore, it imposes the question of how those external and internal technologies work as well as how the workflow of them looks like and, given the recent developments in the field of AI, it is also interesting to look at how these will change the conventional processes of image stabilization extending the research question to about how the future of image stabilization will look like. To supplement the factual research, an experiment was conducted comparing the performance of two commonly used stabilization tools of the non-linear editors “Adobe Premiere Pro” and “Blackmagic Design DaVinici Resolve”. This experiment was expanded by a survey asking participants about their perception of the resulting video clips of the experiment.

2 Method

This thesis is based on research about currently existing methods for image stabilization and methodological proposals for future applications. In order to gain further understanding of how commonly used stabilization methods compare to each other, an experiment with a connected survey was conducted. My methods for all of the contents of this thesis are described below.

2.1 Evaluating existing research

To gain an understanding of the existing research, I searched for articles, books, and papers using the services that were provided to me by the UVU Fulton Library as well as the Library of the University of Applied Sciences St. Pölten. Through them, I was able to access digital content libraries such as IEEE, ACM, SpringerLink, Computer Source, and Sage Online. I also conducted research using Google Scholar, as well as Research Gate. Furthermore, I also gained sources by looking up references listed in encyclopedias or within found sources. The terms I used to search for relevant papers included: “Video stabilization”, “motion compensation”, “digital”, “camera stabilization”, “gimbal systems”, “warping”, “feature tracking”, “real-time”, “artificial intelligence”, “machine learning”, “optical systems”, “homography” and “image stabilization”. Reading through the abstracts and introductions of papers enabled me to sort out results that were non-relevant to my research quickly. I also skimmed through books that turned up in my results, looking for mentions of one of the relevant terms. Sources that made it through this selection were reviewed another time by me in detail in order to gain an understanding of their contents and whether their content applies to my thesis and whether it is outdated already. This led to the exclusion of some sources from my thesis.

2.2 Comparison experiment

The two non-linear video editing programs that were chosen by me for the comparison are “Premiere Pro” by Adobe Inc. and “Davinci Resolve” by Blackmagic Design. The reason for limiting my choice to these two programs was that both of them are established on the market and that I had the resources and licenses in order to conduct this experiment. Originally it was also intended to include Avid’s “Media Composer” in this experiment; however, due to technical difficulties and licensing issues during the time I conducted this experiment, I had to remove it from this comparison. The software versions that were used for this experiment were 14.0.0 (Build 572) for Adobe Premiere Pro and 16.2.0.055 for Blackmagic’s Davinci Resolve, respectively. All the experiments were conducted on an Apple 2017 Retina 5K 27 inch iMac running macOS Catalina 10.15.3 as the operating system. Besides the corresponding video editing software and the diagnostic tools that I required for my experiment and system apps provided by the operating system itself, nothing else was installed in order to avoid the impact of external programs on the results as much as possible. From the hardware side, the machine running the experiment was equipped with an Intel(R) Core(TM) i7-7700K with a clock of 4.20GHz as central processing unit (CPU), 64 GB of random access memory (RAM) provided by four 16GB DDR4 RAM sticks with a speed of 2400Mhz, an AMD Radeon PRO 580 with 8GB of video RAM as graphics processing unit (GPU) and as storage a proprietary 2TB NVM Express SSD that is described within the System report of macOS as “APPLE SSD SM2048L”.

The video clip that was used as the base for this research was a 9 second long extracted handheld recording of fall colored leaves within a tree during sunset that I shot during my stay in the United States. The camera movement within the shot consisted of a wide shot of all the branches and leaves that transitioned over to a closeup of leaves. This footage was originally recorded in 4K (3840x2860) 10bit in HEVC Quicktime format at 60 frames per second with a Fujifilm X-T3 camera. For the actual experiment, I selected these 9 seconds out of the 47 seconds long original clip and exported it as FullHD (1920x1080) MP4 format with h264 as codec and a constant bitrate of 10mbps. This decision was made so that playback would be better supported in the following survey.

The experiment itself consisted of importing the clip into a new project within Adobe Premiere Pro and Blackmagic Design Davinci Resolve, respectively. Afterwards, a digital stabilization effect was applied to the clip. In Premiere Pro, this effect called “Warp Stabilizer” was applied through the Effect Browser. In Davinci

Resolve, this effect “Stabilizer” was applied through the tools menu in the camera panel. In Adobe Premiere Pro, the analysis and calculation of the Stabilization effect are started automatically after application. In Davinci Resolve, it has to be started manually. For this experiment, the default parameters of each stabilization tool were used (see Figures 1 and 2)..

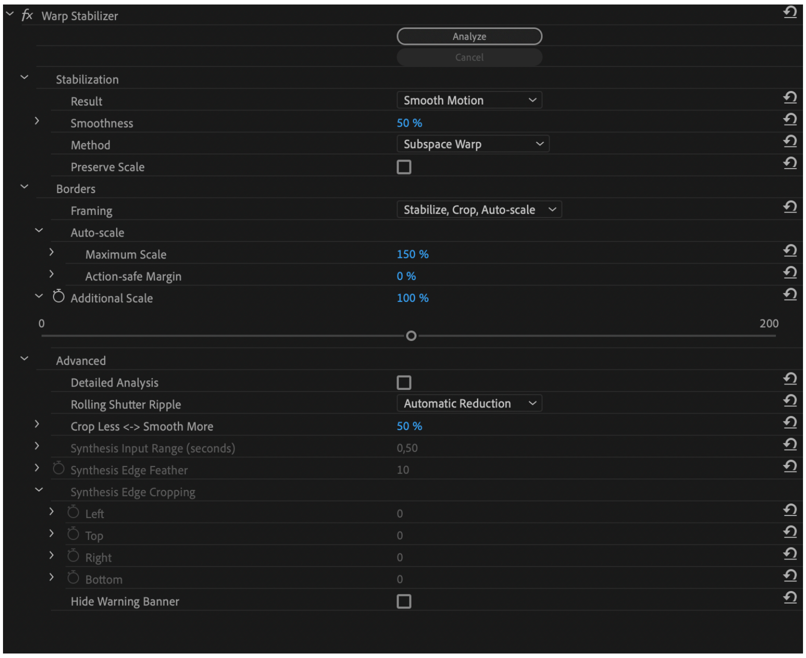

Figure 1: Premiere Pro’s Warp stabilizer default settings used during experiment.

Figure 2: DaVinci Resolve's Stabilizer default settings used during experiment

This process was repeated five times with each program after shutting down and restarting the whole system with each repetition in order to have as consistent conditions as possible. During the experiment, I continuously monitored the duration in seconds needed for the analysis and the following system parameters in order to understand how resource-intensive the stabilization algorithms are. These parameters were: CPU utilization in %, Virtual memory usage in Megabytes, GPU usage in %, and video memory usage in %. In order to have access to and log this data, I was using the activity monitor from macOS, a python script called Syrupy, and an application called iStatMenus. The CPU and memory usage data was pulled directly from the corresponding program process in intervals of 1 second. Due to technical limitations, the GPU usage and memory could only be pulled from the system as a whole and only in intervals of 6 seconds. This is why I started logging the measurements before the actual stabilization took place to get a reference value from which it was possible to see the change in used system resources. The results of this experiment are described in section 4.2.1.

2.3 Survey about experimental results

In order to be able to draw conclusions from the experiment not only from a technical point of view but rather a perceptional as well, a survey was conducted about the resulting video clips from the experiment. The survey was conducted in English and German over the course of a week utilizing the questionnaire service “Typeform” and a web server where the form was embedded. The survey consisted of 3 video clips (one being the source file of the experiment, the other two respectively the resulting clips from the experiment) and seven different questions asking participants about their demographic background as well as whether they were able to notice any differences between the three supplied videos. Furthermore, they were asked to rank these three videos in order of their preference. It is essential to mention that participants were not made aware at any point that the goal of this survey is to rate video stabilization algorithms. Participants were just asked to share their perception of the three video clips. In order to gain participants for the survey, the web address to access was shared via mail and on social media to gain as much reach as possible. The details and results from this survey can be found in section 4.2.2.

3 Analysis

3.1 Image stabilization during capture

Video footage can prove itself challenging to be stabilized after the recording has taken place in post-production. In order to achieve such stabilization, there are different technologies available that have different methods and entry points. There are solutions that achieve stabilization internally in the recording device or attached lens during capture employing functions such as optical image stabilization (OIS) or electronic image stabilization (EIS). In contrast, others achieve stabilized results by engaging externally such as Gimbals or Steadycams. With the rise of processing power and advances in the field of computer science, artificial intelligence has also become a new assisting method contributing to stabilization efforts. (Liang & Shi, 2017)

3.1.1 Internal image stabilization

Internal image stabilization tries to counteract unwanted blurriness on an internal level within the capturing device or the attached lens in order to achieve a certain level of stabilization already during the recording stage. The most common technologies work either on an optical level or electronically effectively stabilizing the recording in realtime. (Sachs et al., 2006)

3.1.1.1 Optical Image Stabilization / Electronic Image Stabilization

There are two approaches commonly used to achieve stabilization on an optical level. lens Shift and Sensor Shift. Otherwise known as "Optical Image Stabilizer" (OIS) and "Electromechanical Stabilization" (EMIS). Both utilize a shift of the optical path in order to counteract movements of the camera system. OIS makes use of a movable lens group or prism by moving them perpendicular to the optical axis in the opposite direction of the movement from the camera system. This is achieved by reading out and processing the live data from two velocity sensors – one being responsible for measurements of the pitch axis, the other one measuring the yaw-axis of the system. This live stream of data is being processed by a processing unit which calculates the needed correction amount and the corresponding direction of the counter-movements. These are forwarded to a control unit that controls an actuator that translates this data into actual movement of the prism or the lens group resulting in a correction of the optical path, so it stays consistent in relation to the image sensor. EMIS works on a very similar level, the main difference being the element being moved.

While OIS achieves the consistency of the optical path utilizing elements within the lens, EMIS moves the imaging sensor in order to stay persistent to the optical path. In comparison to OIS, this provides the advantage that the stabilization is independent of the lens, meaning that stabilization can be achieved regardless of the type of the lens being mounted to the camera. It also uses gyroscopic sensors to measure the movement of the system in combination with a positioning sensor that measures the actual position of the imaging chip at any time. This data is again being processed by a microprocessing unit, which sends directional orders to an actuating mechanism that moves the sensor correspondingly in order to maintain the optical path. (Chen et al., 2007) However, the range of motion that it is able to correct is often reasonably limited – revolving around 1-2 degrees. (Liang & Shi, 2017)

3.1.1.2 Electronical Image Stabilization

Video systems suffer from an additional problem. While single frames can be blurry themselves already as they are captured one by one by an image sensor leading to problems with high-resolution videos, the fact that video footage consists of multiple subsequent frames requires to take into account the shift between these frames. Even if the single frames are not blurry by themselves through the unwanted movement, this shaking can manifest itself as jitter from frame-to-frame over time. As OIS or EMIS is not always enough to prevent this since it prevents the blurriness only on a single-frame level, another system commonly described as "Electronical Image Stabilization" or EIS can be used to counteract this frame-to-frame jitter. Similar to the technologies used by OIS or EMIS systems, a gyroscope is the crucial element to gain a data stream of the movement encountered by the camera device. However, this data is not used to alter the optical path by moving any hardware elements in the camera, but rather on a digital level within the processing software of the imaging chip. The software uses the data from the gyroscope to shift the image correspondingly by the precise number of pixels needed, in order to counteract the jitter that can be present over the course of multiple frames. (Sachs et al., 2006) Nonetheless, this technology on its own also has its limitations though – if the camera jitter becomes strong enough to cause a motion blur within every single frame, shifting the image digitally will not be enough as it leads to distracting sharpness variations. This is why most modern camera models utilize a hybrid system of OIS/EMIS and EIS so that jitter can be counteracted on a single frame level as well as on a frame sequence level. (Liang & Shi, 2017)

3.1.1.3 Artifical Intelligence aided stabilization during capture

As has been mentioned already, capturing smooth and steady footage is a critical principle in videography. This becomes particularly a significant matter for videos being captured with smaller devices, which has become more and more frequent over time with the availability of high-quality cameras being part of smaller systems such as phones. However, these cameras have increasingly limited space to implement smoothing technologies. This is why Artificial Intelligence is gaining a lot of traction in this field as it provides efficient new ways of smoothing images in realtime working together with existing OIS and EIS systems. Similar to other camera systems, phone cameras are particularly susceptible to camera shakes, motion blur, focus breathing, and artifacts such as the “Rolling shutter” distortion. Google’s Software Engineering and Android Camera Team proposed a solution called “Fused Video Stabilization” that claims to address these issues. It utilizes a combination of OIS and EIS during capture divided into three processing stages (see Figure 3). During the first stage described as “motion analysis,” the gyroscope, focus adjustment, and OIS movement data are being extracted at a frequency of 200Hz in order to estimate the rotational movement (roll, pitch, and yaw) of the camera device precisely. During the second stage, described as “motion filtering,” signal processing and machine learning algorithms are applied in order to estimate potential movements and intentions by the person moving the capturing device. This is done by taking the actual motion from the recording device, applying Gaussian filtering on it, and mapping it to a virtual stabilized camera motion. Frames that are being recorded are put in a queue to delay the processing. This gives an ample amount of time to calculate potential future movements and predict the user’s intentions using machine learning. This is done utilizing a proprietary model that has been trained to extract intentional movements from the noise-affected real movement of the capturing device. Furthermore, additional filters are being applied based on the predicted motion – e.g., if the machine learning network predicts an intentional horizontal panning movement, vertical movements would be filtered out more intensively. Finally, during the third “frame synthesis” stage, potential distortions such as the mentioned rolling-shutter and focus breathing are being modeled and removed by applying mesh grids to the input frames and warping each part of the grid separately to counteract the motion and artifacts. The system also takes into account that this processing can lead to bad regions, which occur when the virtual camera path is too stabilized.

The resulting warping creates areas that are out of the original field of view from the original frame. This is predicted in the following frames, and the virtual path is being adjusted to prevent this from happening. Lastly, their solution is also addressing the issue of the sharpness variations, as mentioned earlier, due to motion blurring that can occur withing single frames. It utilizes the masking ability of human perception for motion blur that comes into use when the motion blur follows the same direction as the corresponding motion of the single frame. Using the data from the high-frequency gyroscope and the OIS, they can precisely estimate the amount of motion blur for every single frame. This is achieved by computing the position the camera is pointing at the start and end of the exposure. This is followed by an application of another trained machine learning algorithm that maps the motion blur of past and future frames to the amount of actual camera motion that should be kept and blending this weighted camera motion with the virtual one. (Liang & Shi, 2017)

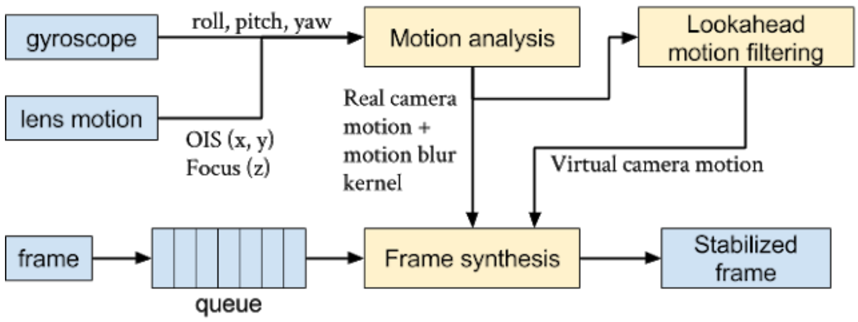

Figure 3: System diagram of Fused Video Stabilization. Image taken from (Liang & Shi, 2017)

3.2 External image stabilization technologies

The downside of a software-based approach in post-production is currently, that in most cases, it inevitably goes hand in hand with a loss of quality or resolution (Chen et al., 2007) and that the algorithms that are being used demand a lot of processing power. The process of “warping” the image, which is one of the standard methods used in commonly utilized software to smoothen the footage, also often leaves noticeable distortions, which are unpleasing to the eye and make the stabilization efforts noticeable. (Zhou et al., 2013) This shows how important it is to avoid random stutters and shaky camera movement in the recording stage as it can result in an enormous amount of effort needed to fix these unwanted movements in the post-production stage. The most commonly used ways of preventing those shaky movements from ending up on the captured footage are to use hardware-based tools such as Steadicams, Glidecams, gimbal-systems, or hand-held rigs. (Grundmann et al., 2011)

All of these systems have in common that they try to counter unwanted movements of the recording device, by balancing out the whole system. However, the way they are achieving this balance is entirely different. (Waldbillig, 2018)

3.2.1 Steadicam

Steadicams and Glidecams mostly rely on counterweights that balance the camera on the axes using physics. Due to their similarities in operation and technology, this thesis will only describe the Steadicam in further detail. The Steadicam is designed to interact with the center of gravity of the recording device. This is done by adding several other masses to the capturing device using a rigid structure (see Figure 4).

Figure 4: Operator using Steadicam rig. Image taken from (Fry, 2016)

This immediately has an effect – this new contraption requires a lot more force to turn angularly, and the center of gravity becomes accessible to the operator, which in turn produces the least angular change – thereby leading to the smallest possible effect of shakiness on the recorded image. Adding masses in two directions leads to a more significant resistance to angular change within all three dimensions meaning that the device will not turn as easy anymore as before. However, given that the carrying of the whole rig and the direction of it is executed by the same muscles, interference is created. This is why Steadicams also make use of a gimbal (see Figure 5). (Holway et al., 2013) However, it is essential to distinguish this gimbal from electric brushless gimbal systems that are described in section 3.2.2.

Figure 5: Three axis gimbal as part of Steadicam structure. Image taken from (Holway et al., 2013)

This three-axis gimbal is located near the new center of gravity of the rig and makes it possible to lift the rig without affecting the structure on an angular level. This gimbal represents another essential part of the Steadicam system as it separates the rather strong lifting force from the small forces that are being applied to the system in order to frame the shot. Furthermore, and what separates the Steadicam from a Glidecam system is that it also utilizes a mechanical arm to provide spatial isolation from the operator and his body movements as well as a specialized vest in order to distribute the often extensive weight of the whole system more evenly on the operator. The arm has a set of components to provide this spatial isolation. Horizontal movements are absorbed by a series of hinges and links that mimic human joints. In contrast, vertical movements are absorbed by two links of the arms that are parallelograms, including internal springs. This has another advantage, as the operator can now move the camera sled both horizontally and vertically without requiring a lot of effort, as well as hold the camera steady when not applying any force at all. (Holway et al., 2013)4/10/20 3:47:00 AM

3.2.2 Electric gimbal systems

Advances in technology have established another external stabilization system on the market, which utilizes gimbals with electronic brushless motors as its main stabilization component. This has made it possible to provide stabilization on a smaller rig, enabling video productions to mount camera systems on different types of vehicles or to carry out shots that would not have been executable before (e.g., handing off the camera from a drone to a walking operator during a continuous take). They are also usually cheaper to rent or purchase than Steadicam systems. However, they also have caveats. They usually have bigger limitations regarding the weight of equipment that is being mounted on to the system. Additionally, introducing electronics to the system can also make things more complicated in case of a failure, as it is not as easy and quick to fix like systems such as the Steadicam, which is solely based on principles of physics. (Waldbillig, 2018) Most of these systems are based on a 3-axis gimbal that addresses tilt, roll, and pan movements, but 4-axis stabilization systems, counteracting unwanted vertical movements additionally, have also been introduced to the market already. (Renée, 2015) For this thesis, I will focus on describing the common 3-axis system. Fundamentally these systems are similar to a Steadicam, as the recording device is mounted to a base plate that is movable within all three of the axes and that it needs to balance out. This means that the electronic gimbal systems also operate based on the laws of physics and the principle of inertia. (Bender, 2015) Gimbal based stabilization systems take advantage of gimbals as a pivoted support system. Three gimbals mounted on to each other on orthogonal pivot axes make it possible for the camera or any other objects, that are mounted on the innermost gimbal to remain independent of the movements that are applied to the external system. ("Gimbal," 2020) Modern electric camera gimbal systems extend this physical concept further by adding three separate brushless servo motors to the system. These motors make it possible to react to any unwanted vibrations that may occur during the capture of video footage precisely. This addition of brushless motors has advantages. In contrast to conventional motors, brushless motors are stepless as they have no transmission via cogwheels, which eliminates vibrations originating from the motor almost entirely. They also make it possible to actively intervene with the position of the recording device in the gimbal system. This enables operators to not only use a gimbal system passively for stabilization but, at the same time, gives them the option to move the camera for desired shots along the axes. (Bender, 2015) Similar to already described stabilization systems, they are controlled by a processing unit that receives its data from sensors, which are described as an "Inertial Measurement Unit" or IMU. (Landrock & Baumgärtel, 2018) This IMU usually consists of gyroscopes, magnetometers, and angular accelerometers. Some units also embed a temperature sensor. (SBG Systems, 2018) The IMU is able to detect the movements of the recording device and reports these motions to the processing unit, which sents its commands to the three brushless servo motors that are positioned in line with the camera lens correspondingly to the axis that they are designated to correct. The brushless servo motor that is mounted on the pitch axis absorbs the unwanted upwards or downwards movement of the camera device, while undesired right or left motion is counteracted by the brushless servo motor on the yaw axis. Lastly, undesired roll motion from one edge to the other is corrected by the brushless servo motor on the roll axis. The addition of this correction on the third axis, the yaw axis, differs these systems from biaxial gimbal systems, which are prone to more unwanted shakiness as there is no absorption of undesired movement from sudden or involuntary turns towards the right or left. Modern gimbal systems need to be balanced out and calibrated by the operator physically and digitally with the supplied software. For instance, it is essential to tell the system what stiffness is required by the motors to handle the weight of the setup correctly. It is also vital to define the maximum angles in order to avoid the gear from accidentally hitting the support structure of the gimbal system, which could lead to damages. Depending on the manufacturer, these options vary, and some even offer unique operating modes that are intended to help the operator setting up and executing shots easier. (Bender, 2015) To process and refine the positional and movement data that is being sent from the IMU, filtering is required in order to make the measurements more accurate. One of the most common filtering methods used is the so-called Kalman filter, which is based on calculations on a variable weighted average ratio. (Canan et al., 2018) It is a set of mathematical equations that provide an efficient and thereby realtime applicable computational solution of the least-squares method. In other words, it makes it possible to estimate past, present, and future states of a system even when the precise nature of the modeled system is unknown. (Welch & Bishop, 1995) Thereby it is possible to estimate current and future values of the IMU system by monitoring the rapidly changing output values of the system, which becomes very useful in image processing, orientation, and motion tracking applications. However, there is research by Canan et al., (2018) that argues that the Kalman filter might be unnecessarily complicated for gimbal stabilization purposes. They propose to use a “Complementary Filter” with a constant weighted average ratio instead. It manages both high pass and low pass filters. While the low pass filters out high- frequency signals that may originate from the accelerometers of the IMU in the case of vibration, the high pass filter eliminates low-frequency signals that may stem from a drift of the gyroscopes of the IMU. The combination of these filters, along with the Complementary one, results in a good signal, while avoiding the complications of a Kalman Filter. Operating wise Gimbal systems are usually either operated in single-op or dual-op mode, with the latter mode enabling a second operator to control the brushless motors directly with remote control, which frees the primary operator from this responsibility and results in him being able to focus on the camera paths and the environment. In this mode, the second operator is usually responsible for the framing of the shot. (Bender, 2015)

3.3 Image stabilization in post-production

Video footage is often being captured without stabilization methods. There are many different reasons why such footage could be captured without stabilization during the recording stage. Some of the technologies that were already presented in this thesis often involve extensive costs and, as such, are often not a viable solution for some video productions. Furthermore, in the case of the Steadicam and the brushless motorized camera gimbal system, skilled operators are needed who are familiar with how these systems work and how to operate them. (Waldbillig, 2018) This is why stabilization technologies are often applied after the recording has taken place during the post-production stage. However, this is not necessarily always the case, and sometimes digital stabilization is also used complementary to stabilization technologies from the recording stage in order to improve and further refine the footage. (Liang & Shi, 2017) For instance, with the introduction of the FX9 camera, Sony showed that combining measurements from a gyroscope during the recording with the processing in post-production can lead to sufficient stabilization results without having to rely on virtual path reconstructions. (Leitner, 2019) Nevertheless, this concept of using gyroscope measurements in order to establish a virtual motion trajectory is not new, and there have been externally mounted solutions on the market (e.g., SteadXP) before that. (Dent, 2018) In order to understand how many of these post-production technologies work, it is important to get to know some of the concepts these algorithms use. Generally speaking, digital video stabilization usually consists of three main processes, which are the matching of feature points, the construction and smoothing of motion trajectories, and finally, the motion compensation. (Souza & Pedrini, 2018; Yoon et al., 2015)

3.3.1 Feature point extraction and image matching

In order to be able to calculate the movement that video recording devices encounter, algorithms are required that make it possible to identify the movement over the course of the consequent frames of the footage. To achieve this, feature points within a single frame need to be defined, extracted, and matched on the next frame. To define these feature points, multiple different methods have varying approaches to detection. One commonly used method for realtime feature detection is called “FAST”. FAST stands for “Features from Accelerated Segment Test”. It is a feature detector commonly used in realtime applications that works by looking at a pixel and the 16 pixels circularly surrounding it. A threshold value is defined, and the pixels are compared to each other. If a set number of the 16 pixels are brighter or darker than the center pixel +/- the threshold, the center pixel is defined as corner and thereby a suitable feature. (Tyagi, 2019)

The movement estimation is also described as “Optical Flow”. So, in other words, Optical flow can be defined as the motion of objects between consecutive frames of the sequence, which is caused by the relative movement between an object that is recorded and the camera. This optical flow can be estimated on a limited base, tracking only “interesting features” within the frame resulting in a “Sparse optical flow” estimation. However, it is also possible to estimate a “Dense optical flow”, by tracking the flow vectors (or movement) of all the pixels in the entire frame (see Figure 6). While Dense optical flow estimations tend to be more accurate, they also have a caveat, being more computationally expensive. (Lin, 2019)

Figure 6: Sparse optical flow estimation (left) vs. dense optical flow estimation (right). Image taken from (Lin, 2019)

3.3.1.1 Homography

From a computer vision perspective, Homography as part of projective geometry represents a motion model that is used to classify motion vectors of a moving image into foreground and background areas. This is relevant for image processing as if one breaks it down to the essentials, what a camera does is mapping the three-dimensional world to a two-dimensional image. In order to be able to let software digitally stabilize an image, it is essential to get the machine to understand what the image is made out of. It is a crucial aspect of computer vision to interpret images three-dimensionally, as based on this data, objects can be measured, and their location within the image can be determined. (Rahmann & Burkhardt, n.d.) This is were Homography comes into play. It is a transformation matrix that can be used to map points in an image to the corresponding points of another image. (Mallick, 2016) One of the common application cases for Homography is to rectify perspective images to a planar one. E.g., mapping perspective aerial photos for a planar map display (see Fig. 7 & 8) (Rossi, 2013). These principles can also be applied to stabilization efforts. Yoon et al., (2015) propose an algorithm that uses classification based Homography estimation for video cameras. Their algorithm goes through four stages, an image segmentation and morphology stage, a feature detection stage, the Homography estimation stage, and finally, a frame warping stage. After defining and extracting the features, they use Homography to detect and match the features in the consequently following frames of the shaky footage.

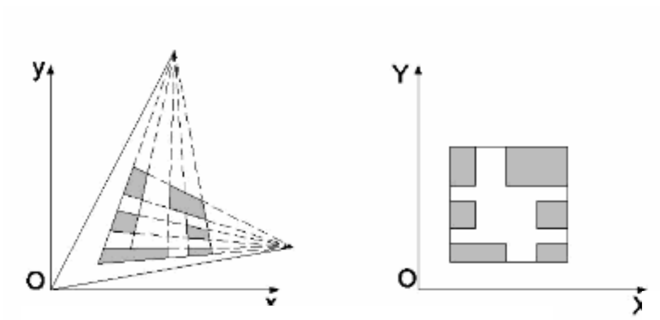

Figure 7: Perspective-to–plane transformation utilizing Homography. Image taken from (Rossi, 2013).

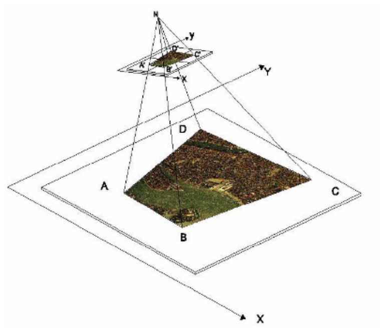

Figure 8: *Mapping of perspective aerial photographic surveys to a planar map. Image taken from (Rossi, 2013)*

3.3.2 Motion estimation and trajectory construction

Marcos R. Souza & Pedrini, (2018) propose the following method for estimating motion and constructing the trajectories. Using the feature points that were extracted, defined, and matched, it possible to construct a motion trajectory over the course of consequent video frames in order to estimate the motion of the camera. To achieve this, a similarity matrix is estimated. It represents the matrix that transforms the features detected in a set frame to the corresponding features in the next frame of the footage. After estimating the similarity matrices for each pair of consequent frames in the video, a trajectory is calculated for each of the factors of the matrices. They consider a vertical, horizontal, rotational, and scaling factor for decomposing the matrices and calculating the resulting trajectories of the factors. These trajectories are then smoothed using a Gaussian filter in order to filter out unwanted oscillations from the calculated camera path (similar to filtering described in section 3.2.2).

3.3.3 Motion compensation

The motion compensation stage adjusts the frames based on the results of the previous stages to compensate for the unwanted movements. There are different methods of how this can be achieved. Avramelos et al., (2018) describe a motion compensation, where the corrected frame is aligned accordingly in the x and y axes using a corrected global motion vector that results from subtracting the difference of the raw and smoothed camera trajectory from the original global motion vector. They also utilize cropping on the corrected frames in reference to the maximum difference between the actual and smoothened camera trajectories in both the x and y-axis in order to avoid black borders that are a result of the translational correction. Yoon et al., (2015) utilize the matched feature points and the smooth camera trajectory to generate warped frames based on the estimation of the homography to compensate for unwanted camera movements.

4 Discussion

After presenting all these technologies and methods, I would now like to draw a comparison between them in order to provide a better picture of what the possibilities are and where some of these mentioned technologies might be limited. I will make this comparison based on my findings during my research as well as my experiment that consisted of a technical comparison of two digital stabilization tools found in two commonly used Non-linear video editing software. It will also be based on the responses of a survey about the results from the experiment that was conducted to gain an understanding about how people perceive the resulting footage from the experiment.

4.1 Stabilization methods used during production

The methods for stabilization during production previously presented were optical image stabilization (OIS), Electromechanical image stabilization (EMIS), electronic image stabilization (EIS), Artificial Intelligence, Steadicam systems, and electronic gimbal systems.

OIS and EIS provide the great advantage that they are usually implemented within the camera or lens and therefore do not require additional external equipment in order to operate. Most common OIS systems are implemented within the lens, which means that in order to take advantage of it, the lens has to support it. EIS, however, works on the processing level and therefore makes the stabilization technology independent from the lens used. The same applies to EMIS as it works on a sensor level. EIS comes at a cost though as this digital processing results in a degradation of the original image. However, EIS also has a distinct advantage over OIS and EMIS concerning video capture. As it is processed digitally, there are no mechanical parts involved, which means that the capture of videos using a high frame rate (in comparison to still photography) is not slowed down by the mechanical parts. The previously mentioned frame-persistent jitter is also not introduced. As was explained before, it can also account for and predict the shift of the next frame.

Additionally, it also has the advantage of being software-based so that it can be continually refined and improved, taking advantage of the advances in fields like Artificial Intelligence. (Knight, 2019)

Some people argue that the introduced electronic gimbal systems could replace Steadicams. (Waldbillig, 2018) This research has shown, however, that this is not the case, as they both still have their distinctive differences and sensible application cases. Steadicams are suitable for bigger camera setups as they are able to handle more weight than gimbal systems, and the weight is being distributed more evenly onto the operator. It also provides 5-axis stabilization, while most gimbal systems are 3-axis based. (Wyndham, 2018) Nonetheless, one of the most significant advantages, is the fact, that it is solely mechanical utilizing principles of physics and does not have any electrical parts that may fail and make a quick repair complicated. (Waldbillig, 2018)

In contrast to the Steadicam, they sometimes also require a second operator in order to precisely frame shots as the operator needs to focus on moving the camera through the location. On the other hand, electrical gimbal systems are cheaper to obtain and also easier to set up, requiring not as much knowledge and experience as the Steadicam does. They also made it possible to do new types of shots as electrical gimbal systems can quickly be passed on to another operator or vehicle. At the same time, a Steadicam is bound to the operator wearing the whole rig. This shows that both systems still have valid purposes. (Waldbillig, 2018)

4.2 Post – production based stabilization methods

As part of this thesis, a comparison experiment was set up to see how two stabilization tools of two commonly used video editing applications compare to each other on a technical as well as perceptional level. This was done by using a reference video clip that was separately stabilized within the two designated programs and comparing technical parameters to evaluate which tool proved to be more efficient and faster. Furthermore, a survey was conducted where participants were asked to answer questions about the resulting video clips in order to gain an understanding of how well humans perceive the results of the stabilization algorithms. The results of this experiment and survey are described below.

4.2.1 Experimental comparison of two post-processing stabilization tools

For this experiment, each program was set to have five runs of applying the stabilization algorithm to the test clip and measuring the performance of it by monitoring and logging system usage parameters. Based on the results, it can be said that the algorithm from DaVinci Resolve seems to be ahead of the algorithm from Premiere Pro. However, the data has shown some interesting insights that are described below and suggest a probable explanation for this significant difference between the programs.

This is the resulting data from the measurements taken during the experiment. Measurement values from during the run were averaged for the run – measurement values from before the runs represent the actual measurement value taken right before the application of the stabilization algorithms:

Table 1: Results of measuring the perfomance of Adobe Premiere Pro's "Warp Stabilizer" tool

Point of measurement	CPU Usage in %	RAM (MB)	GPU usage in %	GPU VRAM %	Duration
Before Run 1	4,70	10423,13	0,00	92,00	
Run 1	246,44	14041,97	0,33	98,00	60,06
Before Run 2	9,20	10277,04	1,00	90,90	
Run 2	241,87	14012,74	0,1	91,53	59,16
Before Run 3	9,4	10149,6	0,00	96,00	
Run 3	241,47	13991,31	0,10	96,97	59,34
Before Run 4	9,30	10425,01	0,00	98,40	
Run 4	246,98	14119,23	0,30	98,67	61,62
Before Run 5	4,9	11419,60	1,00	98,60	
Run 5	247,04	14995,09	0,25	98,97	60,98

Table 2: Results of measuring the performance of Blackmagic Design DaVinci Resolve's "Stabilizer" tool

Point of Measurement	CPU Usage in %	RAM (MB)	GPU usage in %	GPU VRAM %	Duration
Before Run 1	20,20	12702,95	2,00	64,30	
Run 1	110,18	13237,09	27,00	68,20	5,82
Before Run 2	19,50	12782,83	2,00	84,30	
Run 2	96,13	13129,35	23,00	88,30	5,60
Before Run 3	24,60	12761,92	2,00	87,40	
Run 3	106,56	13155,34	26,00	91,40	5,78
Before Run 4	22,90	12215,62	6,00	40,00	
Run 4	106,39	12663,47	31,50	45,90	5,98
Before Run 5	20,80	12331,84	22,50	43,10	
Run 5	103,44	12704,00	43,50	48,50	5,82

The analysis of the duration time for the stabilization algorithms from Premiere Pro and DaVinci Resolve (see Figure 9) has shown that there were significant differences. While the stabilization time stays pretty much consistent within each program during all of the five runs, there was a gap between the programs when one compares them to each other. DaVinci Resolve was, on average, about 12 x times faster than Premiere Pro in executing the stabilization task at hand in this experiment.

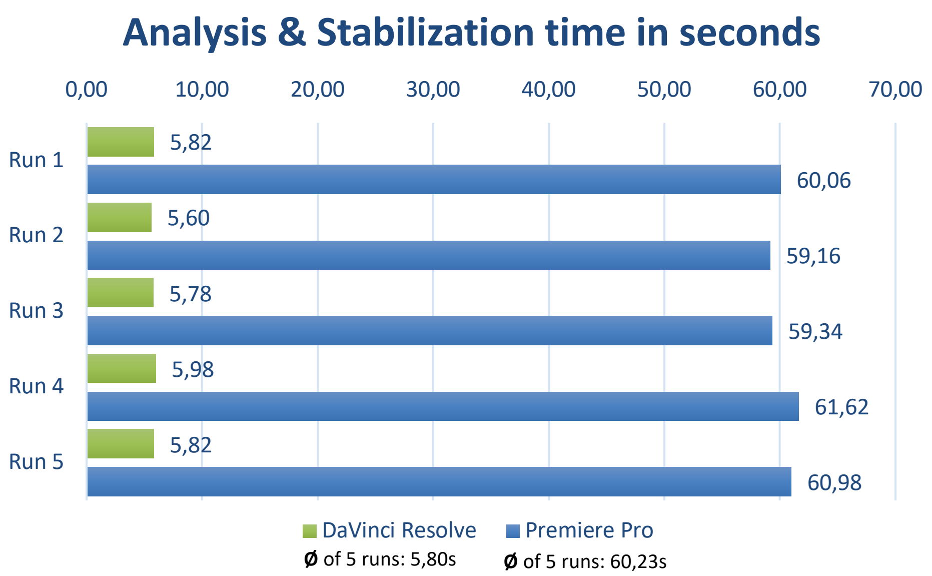

Figure 9: Measurement results for duration of runs during experiment.

Taking a look at the results regarding the average CPU usage (see Figure 10) one can see that the programs utilize the CPU consistently similar over the course of the five runs, however, in comparison to each other the average usage is very different, and this could be the first hint to as why the needed calculation time varies that much. Premiere Pro's Warp stabilizer used on average almost 2,5 x times more processing power than its competitor DaVinci Resolve. It is important to note that these measurements show values above 100%. This is because of the fact that in Multicore processors (the Intel(R) Core(TM) i7-7700K that was used for this experiment has four cores), each core is being measured separately on its own.

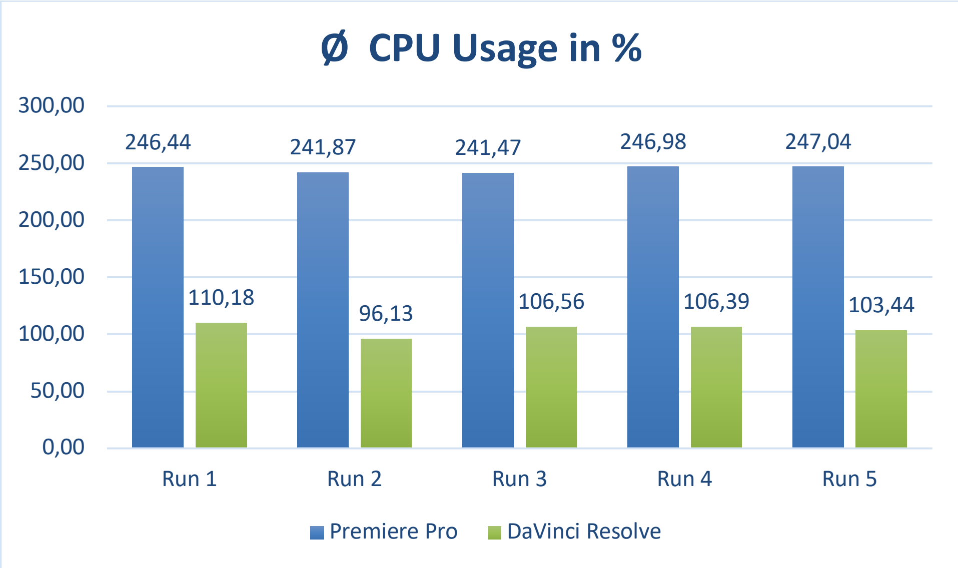

Figure 10: Average CPU usage measured during runs in experiment.

The average RAM usage (see Figure 11) also stays consistent with the previous measurements, as Premiere Pro seems to take up slightly more RAM for its calculations than DaVinci Resolve does. However, the difference is not as significant as in the other parameters. The only measurement that is out of the ordinary is the fifth run of Premiere Pro. For unknown reasons, there was a spike of about 1 GB in comparison to the other runs.

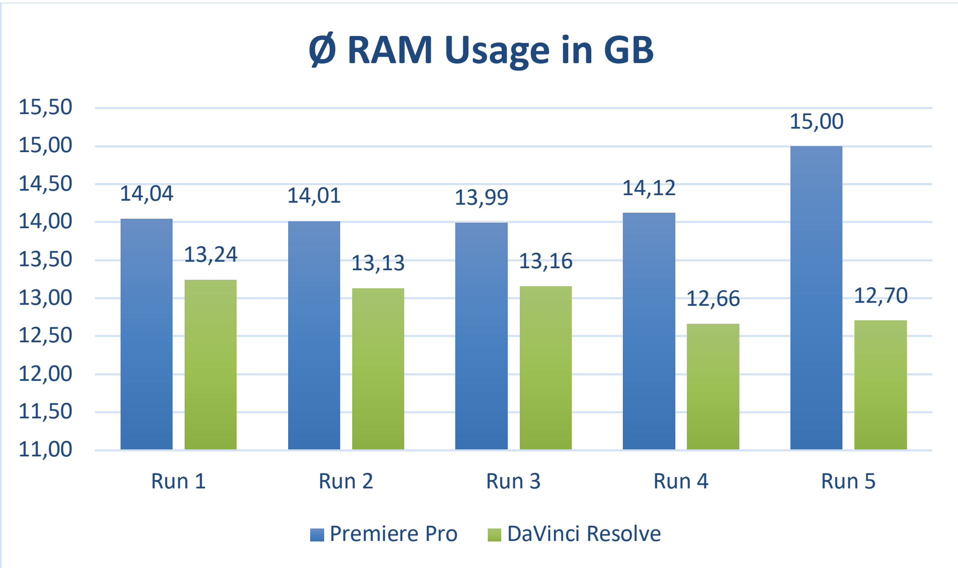

Figure 11: Average RAM usage measured during runs in experiment.

Comparing the GPU usage (see Figure 12), it becomes very apparent how both tools approached their task differently. The values in Figures 12 & 13 represent the GPU usage of the whole system. As it was impossible to measure the usage just in relation to the corresponding process due to technical limitations, a reference value was measured right before the tools were applied to the clip. These values are referenced with “Before Run”. It became apparent that while Premiere Pro’s Warp Stabilizer does not seem to use the GPU at all, DaVinci Resolve does use it extensively. There is a recognizable spike (represented by the blue error bars in Figure 12) of about 23,3% on average in GPU usage during the runs with DaVinci Resolve. Interestingly for Premiere Pro there was an actual drop of usage during Run 2 and 5, however these seem to be negligible and the cause for this was undeterminable. These results suggest that DaVinci Resolve did rely on GPU acceleration in order to calculate the stabilization, while Premiere Pro did not or at least did not make extensive use of it. Why Premiere Pro did not make use of GPU acceleration during the experiment runs is questionable, as according to *Adobe Premiere Pro System Requirements* (2020), the GPU used in the test system, an AMD Radeon Pro 580 with 8GB Video memory, is officially supported for GPU acceleration. The Premiere project was also configured to utilize GPU acceleration, and Warp Stabilizer is listed as an “accelerated Effect” in the official documentation and within the program. (*Types of effects in Premiere Pro*, 2018). It can be assumed that this is also the primary cause of the differences in calculation times (see Figure 9), as outsourcing the calculations from the CPU to the GPU usually results in faster calculations, given that GPU’s are more suitably designed for extensive volume calculations. (White, 2017)

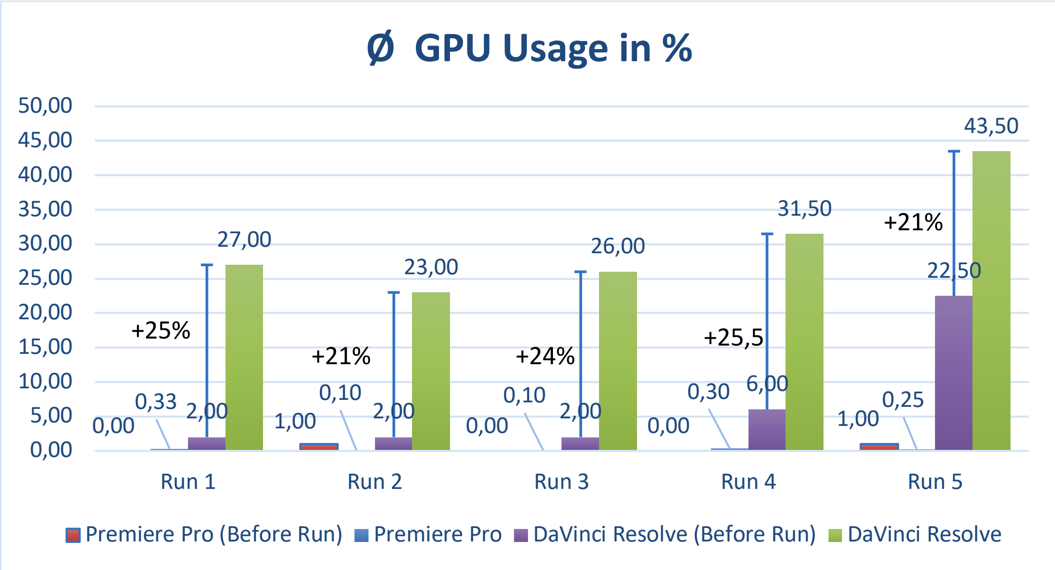

Figure 12: Average GPU usage of system before and during runs of experiment.

Taking a look at the measurements from the average usage of GPU video memory (see Figure 13), one is also able to draw interesting conclusions. First of all, due to the same technical limitations, these values are also measured as the usage of the whole system. In contrast to the regular memory, these values are the usage of the video memory of the GPU (being equipped with 8 Gigabyte of video memory) in percentages and not the actual amount used. Taking a focus on Premiere Pro, one gets values that fall in line with the assumption that it did not really rely on the GPU extensively during these calculations as values do not seem to increase a lot from before the experiment to over the course of the experiment. This applies to all runs with the exception of Run 1 where a significant increase of 6% is noticeable, however, given that these values are measured from the whole system, and this kind of spike is not noticeable in the other runs one can assume that this was caused by something else than the Warp Stabilizer. Shifting attention to the measurements for DaVinci Resolve, we can see reasonably consistent increases during all runs. On average, the increase in usage is about 4,64%.

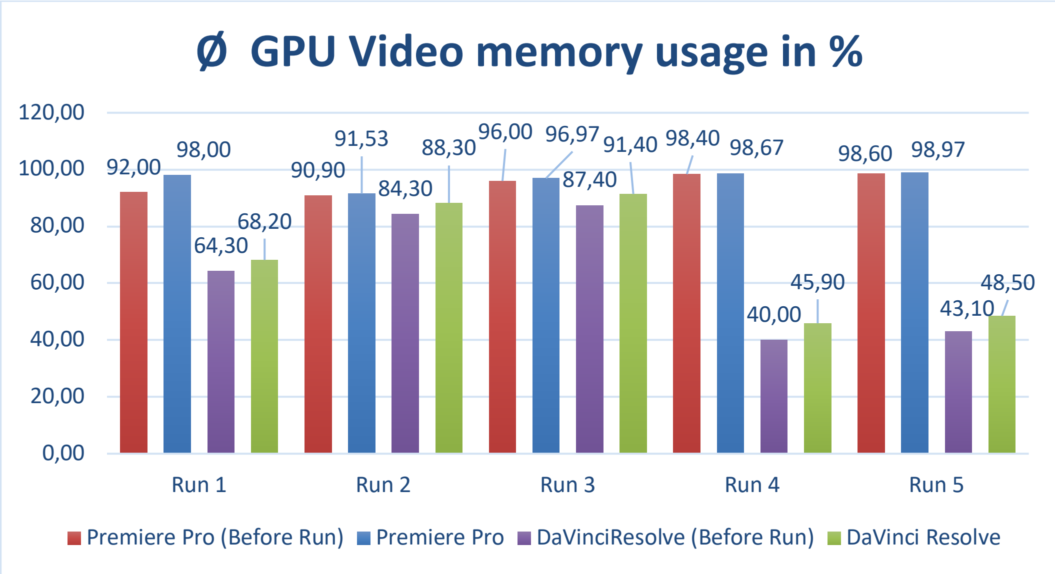

Figure 13: Average GPU video memory usage before and during runs of experiment.

These results have shown that in the experiment that has been carried out, Blackmagic Design's DaVinci Resolve seems to have the upper hand from a technical perspective. It can be assumed that this is mostly due to the utilization of GPU acceleration and that results would look very different if Premiere Pro would have utilized it to the same extent. However, since the technical perspective is not the only component that matters, but also how the results are perceived by humans, a a survey was also carried out to see how participants would rate the resulting video clips.

4.2.2 Survey about experimental results

Following the experiment, a survey was conducted in order to gain insights on how the results of the experiment are perceived by people as it is also relevant that a stabilization algorithm is not only efficient but also delivers enjoyable results. During the one week long period, the survey was conducted 147 people from multiple age groups, and genders participated (see Figure 14 and 15).

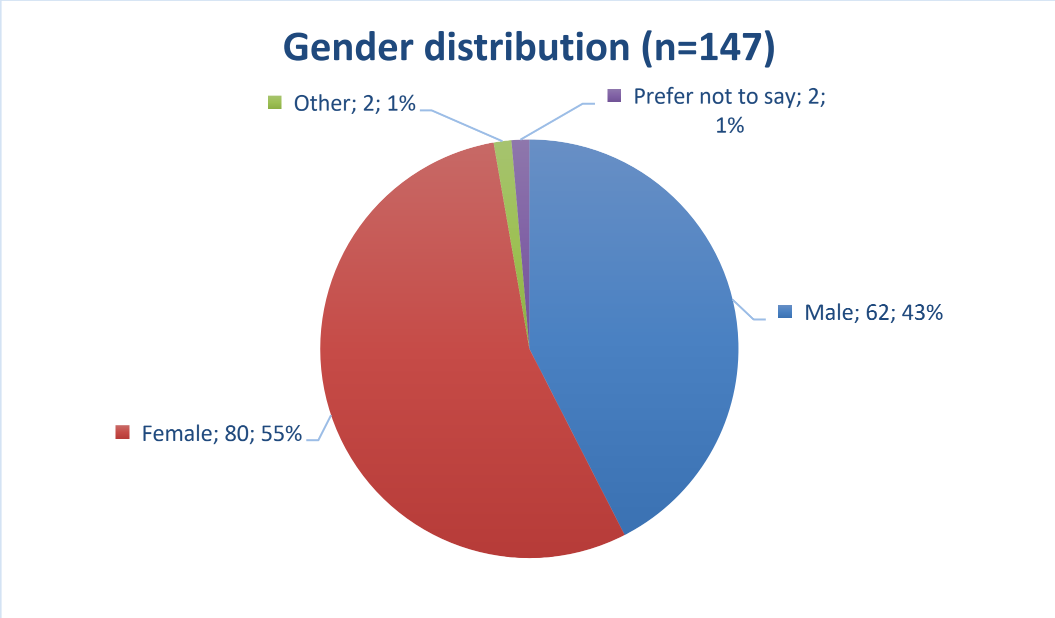

Figure 14: Gender distribution of participants in survey.

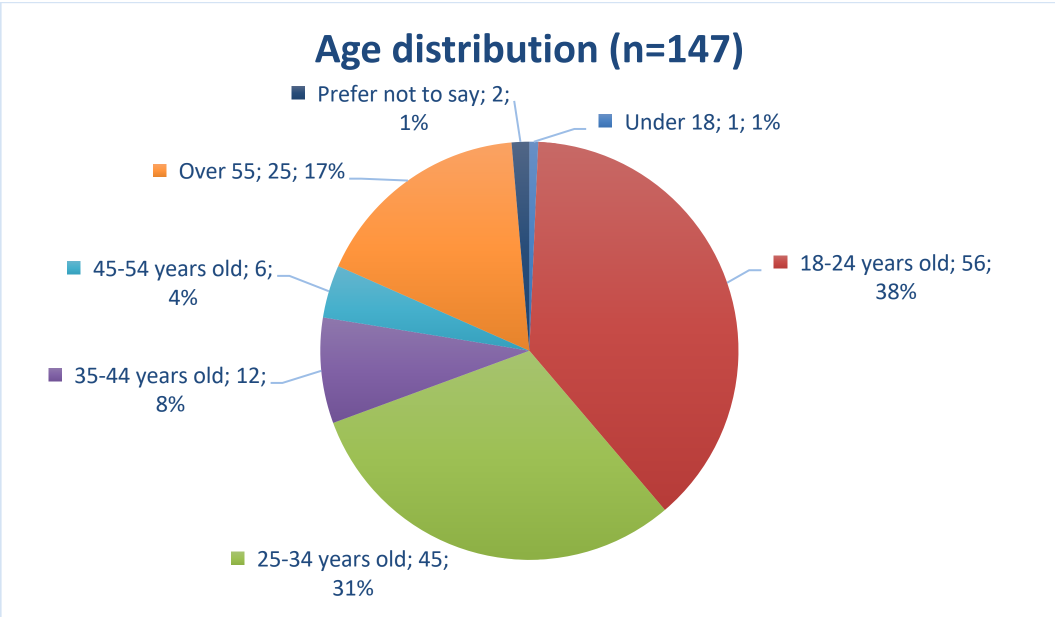

Figure 15: Age distribution of participants in survey.

As described in section 2.3, participants were asked to watch three video clips in a row – one being the source clip of the experiment, the other two being the resulting clips from the experiment. Clip 2 was stabilized in Premiere Pro, clip 3 in Davinci Resolve. The first question was to determine whether participants would recognize any difference at all between the clips. The answers show a very divided opinion (see Figure 16). Across all age groups, the number of participants recognizing a difference is not very significantly bigger than the number of participants not recognizing any difference. However, if one takes a look at the age groups in detail, there seems to be a possible correlation between age and the perception of differences. In the age brackets of 18-34, significantly more participants were able to recognize differences, while in the age brackets above, more participants were unable to see any differences.

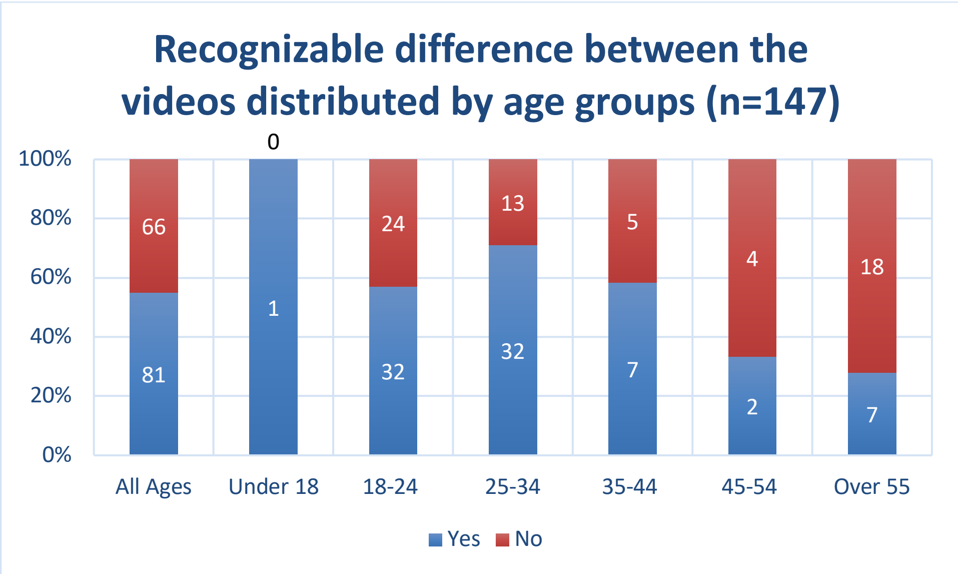

Figure 16: Results of survey distributed by age groups regarding recognition of differences between the clips.

Following this question, participants who recognized a difference were asked to freely describe what differences they had noticed. One aspect that was mentioned significantly often was that people perceived a difference in focus shift between the different clips. It is essential to mention here that the original source clip already included a focus shift that was still present in the experiment clips as well– so it can not be clearly determined whether this perception is based solely on the original focus shift from the source clip or whether some additional focus changes have been introduced through the processing that were recognized by participants. About a third of participants, who were able to recognize something, mentioned that they recognized that some kind of stabilization was applied to at least one of the clips, or that the first clip seems to be shakier than the others. Furthermore, it is important to mention that at no point in the survey participants were made aware of the background of this survey – they were simply asked to freely describe their impressions after watching the videos. About a sixth of answers were also related to the speed of the clips. These persons felt that the clips would differ in playback speed. However, this was not the case. Another aspect that was recognized by some participants was the fact that Clip 2 & 3 seemed to be cropped in, or as one participant put it, “used the zoom function in a different way”. This is one of the typical side-effects of digital stabilization described in section 3.3 as some image resolution is often used by the algorithms to counter shakiness.

Lastly, participants were asked to rank the three video clips in order of their preference. This question was also asked to people who could not recognize any differences, as I assumed that the clips might still have a subconscious impression on participants. The results confirmed my assumptions as 29 of the 66 people, that were unable to recognize any differences were still able to specify a ranking of the clips. The other 37 participants declared that all of the three clips are too similar in order to rank them based on preference. The ranking results of all participants are visualized in Figure 17. One noticeable aspect is that the stabilized clips were perceived, preferably by participants. However, the unstabilized source clip has also left a positive impression on many.

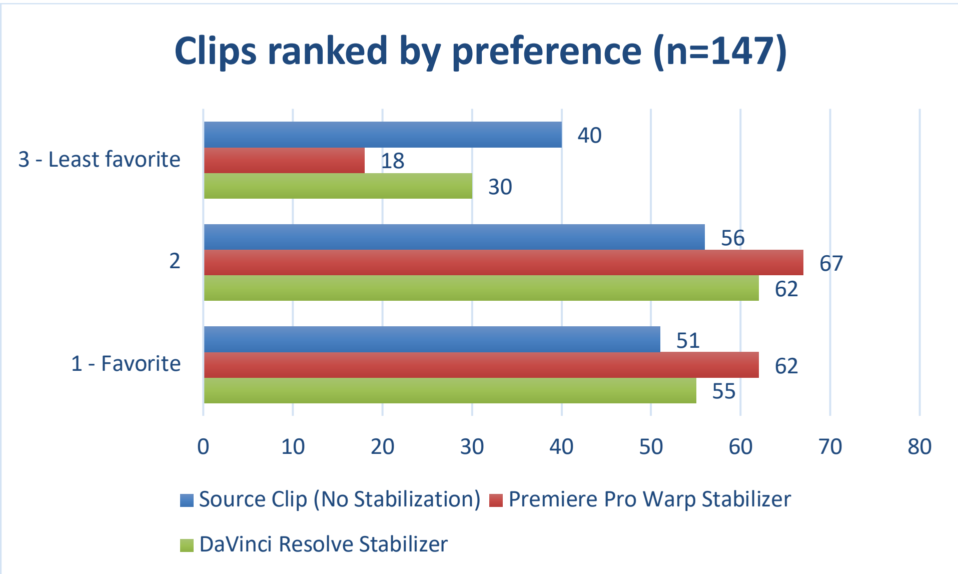

Figure 17: Results of survey regarding ranking of clips by preference.

Refining the results, by excluding all the participants that ranked all clips equally, results in a distribution like in Figure 18. This slightly different distribution further emphasizes that the resulting clip from Premiere Pro was perceived most favorably. The runner up is DaVinci Resolve’s clip followed by the source clip.

Nonetheless, it also shows that the differences in perception are not extremely significant.

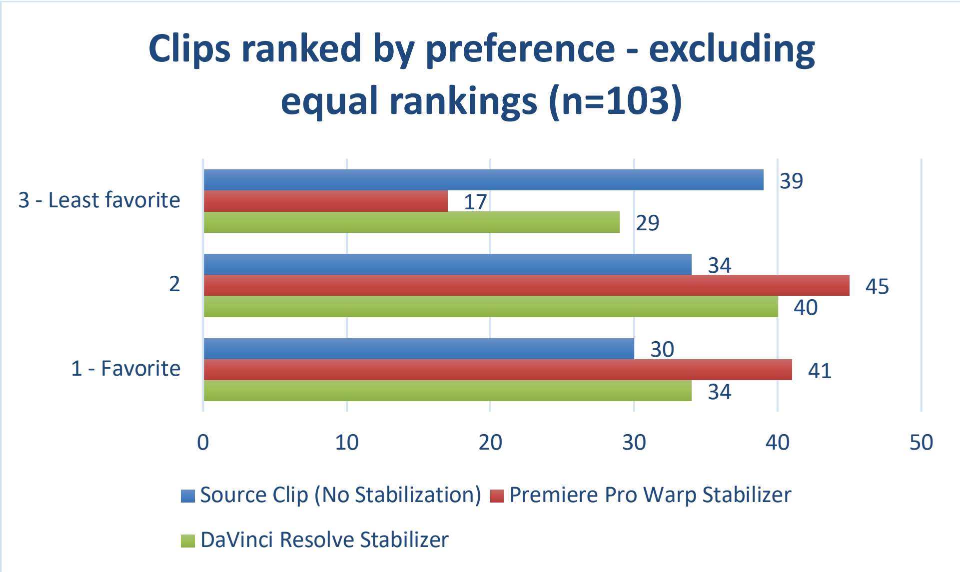

Figure 18: Results of survey regarding ranking of clips by preferences - adjusted by removing answers without ranking.

5 Conclusion

In conclusion, the research conducted for this thesis has shown that there is a multitude of technologies available to mitigate the problem of unstable footage. It is a field that steadily redevelops itself, and new methodologies are discovered frequently. This process has been expedited with the introduction of artificial intelligence, as new algorithms are able to improve upon themselves consistently. It is impressive how digital algorithms are continually getting more effective and efficient, giving us the possibility to fix mistakes that were made during capture in post-production. Additionally, it also made clear that there is no method that „beats them all". All methods have their advantages and disadvantages, as well as their suitable application fields. The experiment I conducted has shown that while there are differences in how commonly used digital stabilization tools process the footage, it does not necessarily mean that these differences are perceivable by persons that well and make up for a big difference. However, this experiment also had flaws that could be improved upon if it was conducted again. For instance, it would make sense to use a wider range of clips for the experiment to have different sceneries to compare from. It would also be useful to have a scene at least, that does not include a change in focus in order to avoid the unclear results described in section 4.2.2. My intention was to have a clip that proves to be more challenging to the algorithms, but this, unfortunately, led to people potentially having misinterpreted the results. From a technical point of view, my technical limitations in the macOS ecosystem also had an impact on the experiment. There are more thorough monitoring solutions available for Windows, and for Nvidia graphic cards. However, as macOS and AMD GPUs are still in use by many in the creative industry, it still has a purpose to conduct this research on these system configurations. The survey could also be expanded further, maybe giving the participants more guidance of what to look out for. This could lead to more precise results.

Literature

Adobe Premiere Pro System Requirements. (2020, March 19). https://helpx.adobe.com/premiere-pro/system-requirements.html#gpu-acceleration

Avramelos, V., Wallendael, G. V., & Lambert, P. (2018). Real-Time Low-Complexity Digital Video Stabilization in the Compressed Domain. *2018 IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, 1–5. https://doi.org/10.1109/ICCE-Berlin.2018.8576211

Bender, M. (2015). *Steadycam vs. Brushless Gimbal: Technik, Handhabung und Ästhetik* [University of Applied Sciences Mittweida]. https://monami.hs-mittweida.de/frontdoor/index/index/year/2015/docId/5145

Canan, S., Dere, E., & Ozcan, M. (2018, August). *(PDF) Three Axis Gimbal Design and Its Application*. ResearchGate. https://www.researchgate.net/publication/326995081_Three_Axis_Gimbal_Design_and_Its_Application

Chen, B.-Y., Lin, J.-S., & Huang, W.-T. (2007). Stabilizing Video While Keeping Resolution and Capturing Intention. *ACM SIGGRAPH 2007 Sketches*. https://doi.org/10.1145/1278780.1278886

Dent, S. (2018, January 24). *SteadXP's DSLR stabilizer is a gimbal with no moving parts*. Engadget. https://www.engadget.com/2018-01-24-steadxp-s-dslr-stabilizer-impressions.html

Gimbal. (2020). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Gimbal&oldid=946944913

Golik, B. (2006). *Development of a Test Method for Image Stabilizing Systems*. 48.

Grundmann, M., Kwatra, V., & Essa, I. (2011). Auto-directed video stabilization with robust L1 optimal camera paths. *CVPR 2011*, 225–232. https://doi.org/10.1109/CVPR.2011.5995525

Holway, J., Hayball, L., & Safari, an O. M. C. (2013). *The Steadicam® Operator's Handbook*. https://www.safaribooksonline.com/complete/auth0oauth2/&state=/library/view//9780240811659/?ar

Hsu, Y., Chou, C., & Shih, M. (2012). Moving camera video stabilization using homography consistency. *2012 19th IEEE International Conference on Image Processing*, 2761–2764. https://doi.org/10.1109/ICIP.2012.6467471

Knight, J. (2019, May 4). *Why EIS Is Actually Better Than OIS for Videos*. Gadget Hacks. https://smartphones.gadgethacks.com/news/why-eis-is-actually-better-than-ois-for-videos-0195575/

Landrock, H., & Baumgärtel, A. (2018). *Die Industriedrohne -- der fliegende Roboter: Professionelle Drohnen und ihre Anwendung in der Industrie 4.0*. https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1823520

Leitner, N. (2019, September 13). Sony FX9 Announced – Full-Frame, Fast Hybrid Autofocus, Dual ISO Camera. *Cinema5D*. https://www.cinema5d.com/sony-fx9-announced-full-frame-fast-hybrid-autofocus-dual-iso-camera/

Liang, C.-K., & Shi, F. (2017, October 11). Fused Video Stabilization on the Pixel 2 and Pixel 2 XL. *Google AI Blog*. http://ai.googleblog.com/2017/11/fused-video-stabilization-on-pixel-2.html

Lin, C. (2019, April 24). *Introduction to Motion Estimation with Optical Flow*. AI & Machine Learning Blog. https://nanonets.com/blog/optical-flow/

Liu, F., Gleicher, M., Jin, H., & Agarwala, A. (2009). Content-preserving warps for 3D video stabilization. *ACM Transactions on Graphics*, *28*(3), 1. https://doi.org/10.1145/1531326.1531350

Mallick, S. (2016, January 3). *Homography Examples using OpenCV (Python / C ++) | Learn OpenCV*. https://www.learnopencv.com/homography-examples-using-opencv-python-c/

Rahmann, S., & Burkhardt, H. (n.d.). *Praktikumsversuch Kamerakalibrierung*.

Renée, V. (2015, September 26). *Jockey Motion is a 4-Axis Gimbal That Can Also Upgrade Your 3-Axis Ronin & AllSteady*. No Film School. https://nofilmschool.com/2015/09/jockey-motion-4-axis-gimbal-also-upgrade-3-axis-ronin-allsteady

Rossi, G. (2013). *The Homography transformation*. http://www.corrmap.com/features/homography_transformation.php

Sachs, D., Nasiri, S., & Goehl, D. (2006). *Image Stabilization Technology Overview*.

SBG Systems. (2018). *Inertial Measurement Unit for Gimbal Camera Stabilization*. SBG Systems. https://www.sbg-systems.com/wp-content/uploads/Ellipse_2_Micro_Series_Leaflet.pdf

Smaga, S. (2003). Tremor. *American Family Physician*, *68*(8), 1545–1552. https://www.aafp.org/afp/2003/1015/p1545.html

Souza, Marcos Roberto e, & Pedrini, H. (2018). Combination of local feature detection methods for digital video stabilization. *Signal, Image and Video Processing*, *12*(8), 1513–1521. https://doi.org/10.1007/s11760-018-1307-8

Souza, Marcos R., & Pedrini, H. (2018). Digital video stabilization based on adaptive camera trajectory smoothing. *EURASIP Journal on Image and Video Processing*, *2018*(1), 37. https://doi.org/10.1186/s13640-018-0277-7

Tyagi, D. (2019, January 2). *Introduction to FAST (Features from Accelerated Segment Test)*. Medium. https://medium.com/data-breach/introduction-to-fast-features-from-accelerated-segment-test-4ed33dde6d65

Types of effects in Premiere Pro. (2018, July 20). https://helpx.adobe.com/premiere-pro/using/effects.html#ListofAccelerated32bitandYUVeffectsinPremierePro

Waldbillig, O. (2018, January). *Article*. Oliver Waldbillig - Steadicam Operator. http://www.oliver-waldbillig.lu/steadicam-or-gimbal/

Welch, G., & Bishop, G. (1995). *An introduction to the Kalman filter*.

White, J. (2017, March 31). *The advantages of GPU acceleration in computational finance*. Elsen. https://elsen.co/blog/the-advantages-of-gpu-acceleration-in-computational-finance/

Yoon, I., Jeon, S., Jeong, S., & Paik, J. (2015). Video stabilization using classification-based homography estimation for consumer video camera. *2015 IEEE 5th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, 116–117. https://doi.org/10.1109/ICCE-Berlin.2015.7391209

Zhang, L., Xu, Q.-K., & Huang, H. (2017). A Global Approach to Fast Video Stabilization. *IEEE Transactions on Circuits and Systems for Video Technology*, *27*(2), 225–235. https://doi.org/10.1109/TCSVT.2015.2501941

Zhou, Z., Jin, H., & Ma, Y. (2013). Plane-Based Content Preserving Warps for Video Stabilization. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2299–2306. https://doi.org/10.1109/CVPR.2013.298

List of Figures

FIGURE 1: PREMIERE PRO'S WARP STABILIZER DEFAULT SETTINGS USED DURING EXPERIMENT. 5
FIGURE 2: DAVINCI RESOLVE'S STABILIZER DEFAULT SETTINGS USED DURING EXPERIMENT 5
FIGURE 3: SYSTEM DIAGRAM OF FUSED VIDEO STABILIZATION. IMAGE TAKEN FROM (LIANG & SHI, 2017) 10
FIGURE 4: OPERATOR USING STEADICAM RIG. IMAGE TAKEN FROM (FRY, 2016) 12
FIGURE 5: THREE AXIS GIMBAL AS PART OF STEADICAM STRUCTURE. IMAGE TAKEN FROM (HOLWAY ET AL., 2013) 13
FIGURE 6: SPARSE OPTICAL FLOW ESTIMATION (LEFT) VS. DENSE OPTICAL FLOW ESTIMATION (RIGHT). IMAGE TAKEN FROM (LIN, 2019) 17
FIGURE 7: PERSPECTIVE-TO–PLANE TRANSFORMATION UTILIZING HOMOGRAPHY. IMAGE TAKEN FROM (ROSSI, 2013). 18
FIGURE 8: MAPPING OF PERSPECTIVE AERIAL PHOTOGRAPHIC SURVEYS TO A PLANAR MAP. IMAGE TAKEN FROM (ROSSI, 2013) 18
FIGURE 9: MEASUREMENT RESULTS FOR DURATION OF RUNS DURING EXPERIMENT. 24
FIGURE 10: AVERAGE CPU USAGE MEASURED DURING RUNS IN EXPERIMENT. 25
FIGURE 11: AVERAGE RAM USAGE MEASURED DURING RUNS IN EXPERIMENT. 25
FIGURE 12: AVERAGE GPU USAGE OF SYSTEM BEFORE AND DURING RUNS OF EXPERIMENT. 27
FIGURE 13: AVERAGE GPU VIDEO MEMORY USAGE BEFORE AND DURING RUNS OF EXPERIMENT. 28
FIGURE 14: GENDER DISTRIBUTION OF PARTICIPANTS IN SURVEY. 29
FIGURE 15: AGE DISTRIBUTION OF PARTICIPANTS IN SURVEY. 29
FIGURE 16: RESULTS OF SURVEY DISTRIBUTED BY AGE GROUPS REGARDING RECOGNITION OF DIFFERENCES BETWEEN THE CLIPS. 30
FIGURE 17: RESULTS OF SURVEY REGARDING RANKING OF CLIPS BY PREFERENCE. 31
FIGURE 18: RESULTS OF SURVEY REGARDING RANKING OF CLIPS BY PREFERENCES - ADJUSTED BY REMOVING ANSWERS WITHOUT RANKING. 32

List of Tables

TABLE 1: RESULTS OF MEASURING THE PERFOMANCE OF ADOBE PREMIERE PRO'S "WARP STABILIZER" TOOL 22
TABLE 2: RESULTS OF MEASURING THE PERFORMANCE OF BLACKMAGIC DESIGN DAVINCI RESOLVE'S "STABILIZER" TOOL 23