Tampere University

Leevi Uosukainen

EVENT CAMERAS FOR MOBILE IMAGING

Handshake blur removal and the technology life cycle

Master of Science Thesis
Faculty of Engineering and Natural Sciences
Examiners: Prof. Kari Koskinen, Prof. Joni Kämäräinen, D.Sc. (Tech) Radu Ciprian Bilcu
December 2021

ABSTRACT

Leevi Uosukainen: Event Cameras for Mobile Imaging
Master of Science Thesis
Tampere University
Master of Science (Technology)
December 2021

Event cameras are novel imaging sensors used to capture illumination changes in a scene rather than exposing the pixels to all incoming light for a given time. Together with RGB imaging sensors, they can be used for several image and video enhancement applications. In this thesis it was tested whether it is possible to reduce handshake blur by utilizing event data. It was found out that handshake blur removal is possible. Technology life cycle analysis was conducted as well based on patent data, and it was determined that most likely event camera technology is ongoing growth, the second stage of the life cycle. Evidence of event camera integration towards mobile phones was obtained by examining patent documents related to event camera technology, and some signs referring to the possible future integration was discovered.

Keywords: event camera, handshake deblurring, technology life cycle, patent analysis

TIIVISTELMÄ

Leevi Uosukainen: Event Cameras for Mobile Imaging
Diplomityö
Tampereen yliopisto
Konetekniikan DI-ohjelma
Joulukuu 2021

Tapahtumakamerat ovat kuvantamissensoreita, joiden toimintaperiaate perustuu kuvattavassa kohteessa tapahtuvien kirkkauden muutosten havainnointiin, toisin kuin perinteisissä RGB-sensoreissa joissa pikseleitä valotetaan kohteesta tulevalla valolla ennalta määritetyn ajan. Yhdessä perinteisten sensoreiden kanssa tapahtumakameroita voidaan käyttää useissa eri kuvien ja videoiden ehostamissovelluksissa. Tässä työssä tarkasteltiin mahdollisuutta poistaa valotuksen aikana sensorin liikkeistä aiheutuvaa liikasumeutta RGB-kuvista tapahtumakameran avulla. Kokeissa havaittiin, että sumeuden poistaminen on mahdollista. Lisäksi tarkasteltiin tapahtumakamerateknologian elinkaarta teknologian elinkaarianalyysin avulla, jossa käytettiin lähdeaineistona teknologiaan liittyviä patentteja. Aineiston perusteella pääteltiin, että tapahtumakamerateknologia on mitä ilmeisimmin teknologian elinkaaren toisessa- eli kasvuvaiheessa. Aineiston perusteella pyrittiin myös arvioimaan, kuinka todennäköisesti tapahtumakamerateknologiaa tullaan integroimaan älypuhelimiin, ja siihen viittaavia merkkejä löytyi.

Avainsanat: tapahtumakamera, sumeudenpoisto, teknologian elinkaari, patenttianalyysi

PREFACE

I want to thank all my colleagues who have provided me with insights and helped to accumulate knowledge related to the topics of this thesis. This work could not have been completed without the support from Huawei Technologies Oy (Finland) Co. Ltd and my thesis supervisors, who helped enormously in formulating the topics and finding adequate methods on how to approach them, for which I want to express special gratitude to them. Finally, I want to thank all my friends and family members who have always helped and encouraged me in many positive ways throughout the years of my studies.

In Tampere, 7th December 2021

Leevi Uosukainen

CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

AER	Address-event representation
AR	Augmented Reality
ATIS	Asynchronous time-based imaging sensor
CMOS	Complementary metal oxide semiconductor
DAVIS	Dynamic and active pixel vision sensor
DVS	Dynamic Vision Sesor
FPS	Frames per second
GIF	Guided image filtering
IPC	International Patent Classification
MSE	Mean squared error
PLC	Product life cycle
PSNR	Peak signal-to-noise ratio
RGB	Red, green & blue
SME	Sum of modulus error
SSIM	Structural similarity
tf	Term frequency
tf-idf	Term frequency-inverse document frequency
TLC	Technology life cycle
TOF	Time-of-flight
USPTO	United States Patent and Trademark Office
VR	Virtual Reality

1. INTRODUCTION

The pace of development of mobile imaging in the past decades has been swift, and the era of smartphones has brought high-quality mobile cameras into the hands of millions of people around the world. In the 2010s, the number of smartphones sold per year more than quintupled as units sold per year increased from 296.65 million in 2010 to 1 540.66 million in 2019 [1]. The amount of smartphone imaging sensors manufactured and sold was multiplied by even higher number, since the number of different sensors per smartphone has been increasing as well. This can be observed by looking at the trends on the smartphone market in the United States and compare it to the smartphone and tablet camera module market trends in North America from 2014 to 2020. The data shows that the market size of smartphones grew by 161.7 % [2], when the segment of the camera module market that consists of smartphone and tablet applications grew by 326.9 % at the same time in North America [3]. Trends of the markets in United States and the whole of North America can be assumed to follow similar path due to the size of the United States economy compared to the other nations in the region. In the same time frame, the growth for average price of smartphones in the United States was 18.9 % for consumer devices and 16.2 % for enterprise devices [4], which means that the larger growth of camera module market compared to the smartphone market can not be explained by decreasing smartphone prices. Moreover, the total number of tablet devices shipped, which were included in the same segment as smartphones in the camera module market size data, increased by a mere 5.02 % from 2014 to 2020 [5].

First glimpses of the phenomenon of increased number of camera modules integrated in mobile phones were some dual-camera smartphones in the early 2010s, which had two similar red, green & blue (RGB) color sensors, only with different resolutions. After that, zoom-dedicated, ultra-wide and monochrome sensors made their first appearance in the mobile imaging market, and after that we have seen time-of-flight (TOF) and structured light sensors appear as well. If the trend of new sensors coming to mobile devices continues, event sensor could be one possible candidate for the category. However, major benefits of the integration must be demonstrated before integrating a whole new type of sensor into a mobile device that is produced in large scale. A new type of sensor could either provide new applications by itself, or via synergetic manner together with the other sensors included in the device.

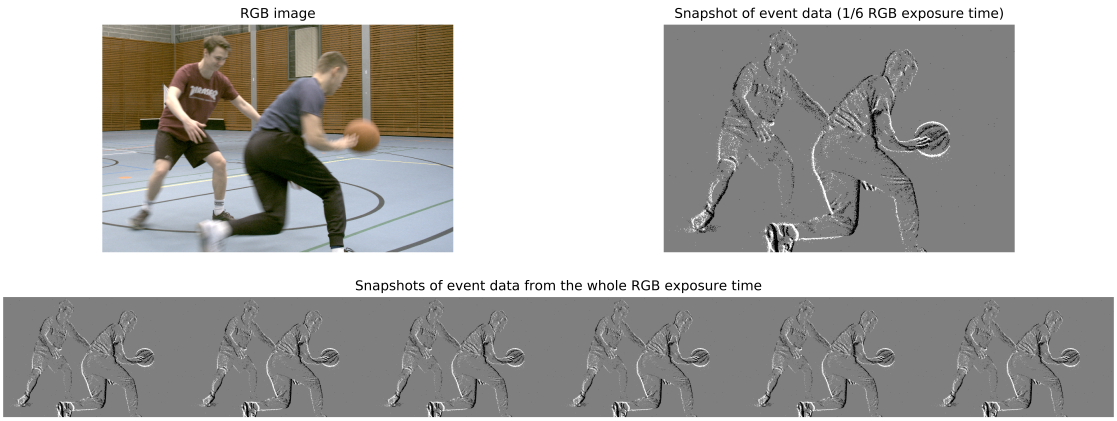

Figure 1.1. *RGB image and corresponding event camera data. Top right image is a single snapshot from the event data, whereas bottom row shows six snapshots, with each corresponding to 1/6 of the RGB image exposure time.*

Event cameras are image sensors where pixels do not gather data by traditional method of exposing themselves to a light source for a given exposure time, but rather continuously and in asynchronous manner, measuring the changes of illumination that a given pixel receives. A sample pair of data captured by RGB and event sensors is displayed in Figure 1.1. As a standalone image sensor in a setup, event cameras might have vast amounts of potential usages in the areas such as machine vision, mobile imaging, augmented reality, security and others. Due to the reasons presented in the previous paragraph, it should be examined what type of advantages a mobile imaging system equipped with an event sensor could possibly offer. Lots of research on event cameras have been conducted in recent years. It is however unclear how much commercial potential the technology actually contains, and hence that should be examined.

Modern mobile cameras still have some shortcomings in the areas of dynamic range and temporal resolution, meaning they are not capable of creating high quality outputs of scenarios where lighting conditions change rapidly between dark and light, and when there is fast movement in the scene during exposure, which are exactly the scenarios where event cameras have significant performative advantage compared to traditional imaging sensors. However, the performance increase in these areas that could be achieved with event camera, could possibly also be achieved via other hardware- or software-based solutions with smaller costs. Algorithms play significant role in the increasing the quality of images captured by contemporary mobile devices, and the costs of mass-producing an algorithm are nonexistent compared to the costs of a large-scale implementation of a new hardware solution. For this reason, the benefits of utilizing event camera for tackling any image quality enhancement problem should be considerably higher than those of any software-based alternative solution.

In this thesis two types of analysis are conducted. First, it will be examined whether it is possible to utilize event data for eliminating handshake blur, which is a type of blur that impairs the image quality and is caused by the movements of imaging sensor during exposure. This will be tested by utilizing a neural network architecture that has been developed for addressing similar problem in the case where only RGB images are used, and by modifying the network to work with event data. The goal of this analysis is to find out how well the handshake blur can be eliminated when event data is utilized. This type of analysis has been chosen due to the fact that mobile devices experience some amount of handshake every time they are used for image capturing, if they are not mounted into stationary setups.

Second analysis that will be conducted is the examining of technology life-cycle of the event camera technology. Motivation behind this analysis is to be able to assess what stage of the technology life cycle the event camera technology is currently experiencing, and to make forecasts about future development. Technology life cycle analysis is done by using patent documents related to event camera technology as the research material, since patents can be considered as proofs of theoretical commercial applicability, which is required in order for them to be granted. Patent data is also available earlier in time compared to other indicators of successful innovation, such as the products that come to the market. Several indicators that can be derived from the data points in the patent documents are used to assess the questions about the current state and future expectations. The patent data gathered for the analysis of technology life cycles is also examined in order to find out which technology sectors are the ones that engage most in research and development activities and which sectors will most likely utilize event camera technology in the future, and how likely it is that mobile phones are among them. The analysis about possible mobile phone integration of event cameras is therefore examined from the technology evolution and innovation perspective, even though it is noted that financial variables such as the costs of a single event sensor module when producing them in large scale also play significant role.

2. EVENT CAMERAS

Event-based sensors are types of *complementary metal oxide semiconductor* (CMOS) imaging sensors. They have completely different working principles compared to traditional RGB imaging sensors that are based on Bayer arrays, and which are used in most imaging applications. Whereas with RGB imaging sensors the working principle is based upon exposing all of the pixels in the sensor to the light received from the scene for a given amount of time, event sensor does not have exposure time at all. Pixels on the event sensor only send information forward when the intensity of the light that the pixel receives goes through a change that is deemed sufficient enough to trigger an event, giving event sensors an asynchronous nature. Most of the imaging sensors tend to be synchronous, which results in information loss during the time that pixels are not exposing themselves to the light coming from the scene, and quality losses in cases where the subject in front of the sensor or the sensor itself is moving during the exposure time. With traditional sensors, information about static scenes is also forwarded along with the parts of the scene that experience change, meaning significant bandwidth and storage will be allocated on transferring and storing information that might be useless, although the storage problem can be addressed with different compressing solutions and the bandwidth problem via static scene detection algorithms. These problems are not apparent with event-based sensors, since events are only passed forward from the pixels which detect a change exceeding a certain threshold in the scene, meaning that if nothing happens in front of the sensor, no data is passed forward.

The building of the first prototype of an event-based sensor was started in the late 1980's and it was completed in 1992 [6]. After 30 years of development, event-based sensors have still not become widely used in any particular industry or segment of consumer products. The process of development continues with a wide range of possibilities across many different sectors, including mobile, robotics, autonomous vehicles, medicine and security.

2.1 Biological inspiration

Event-based sensors are often called silicon-based retinas because of their ability to capture changes of illumination similarly to that of the human eye, asynchronously and with

very low temporal resolution, down to a few microseconds, and even in dark conditions. Traditional imaging sensors require external signals and triggers in order for the image acquisition to take place, whereas in event sensors the events happening in the scene are what causes it to acquire data and pass it forward. It is true that traditional sensors can also acquire data in a manner that seems rather autonomous, one such example being the camera that captures an image when a motion sensor to attached to the system sends a trigger telling the system to capture data, but the underlying principle of the camera needing an external signal in order to start exposing it's pixels to the light coming from the scene is still present.

Similarly to human eye, event sensors do not send information forward if nothing happens on the scene. The reason behind why biological systems behave this way is efficiency, which is a result of biological vision systems undergoing hundreds of millions of years of evolution. animal brains would be overwhelmed with the incoming information if all of the information received from the observed scene would be sent forward for processing. In the human visual system, forwarding only relevant information to the brain makes it possible to reduce the bandwidth from approximately 36 gigabits per second for the raw input to a mere 20 megabits per second for the information forwarded towards brains [7], a decrease of 180 000 %.

2.2 Sensors

There exists several different types of event sensors. The three most dominant types are the dynamic vision sensor (DVS) [8], the dynamic and active pixel vision sensor (DAVIS) [9]. Other types of sensors also exist, such as the asynchronous time-based image sensor (ATIS) [10]. The main shortcoming regarding DVS compared to ATIS and DAVIS is that it does not output any absolute or baseline value of the illumination, only relative changes. After the initial development of DVS, it was found out that more complex applications often require also the baseline illumination detection in order to succeed. Major difference between DVS and DAVIS is also that DAVIS is capable of providing synchronous grayscale image frames together with the event data. In 2021, first sensors which are capable of simultaneous RGB and event capturing were published [11]. The benefit of these types of sensors compared to solely event-based sensors is that integrating them to a system does not require additional space compared to a two-sensor setup.

From the perspective of possible mobile phone integration, small size is an absolute requirement that the sensor must fulfill, even more so if the camera is placed on the front side of the device. Other characteristics that are essential are high dynamic range and high resolution, which are required in order to justify the extra costs that implementing a novel sensor to the devices would result in. It is therefore reasonable to examine and

Sensor description	Sensor size	Dynamic range	Resolution
Asynchronous Temporal Contrast Vision Sensor by Lichtsteiner et al. (2008) [8]	$6 \times 6.3 \quad mm^2$	120 dB	128×128
QVGA Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS by Posch et al. (2011) [10]	$9.9 \times 8.2 \quad mm^2$	143 dB	304×240
Global Shutter Spatiotemporal Vision Sensor by Brandli et al. (2014) [9]	$5 \times 5 \quad mm^2$	130 dB	240×180
Dynamic vision sensor by Samsung (2017) [12]	$8 \times 5.8 \quad mm^2$	> 80 dB	640×480
Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor by Prophesee & Sony (2020) [13]	$4.86 \times 4.86 \quad mm^2$	> 124 dB	1280×720
Dynamic vision sensor by Samsung (2020) [14]	$8.37 \times 7.64 \quad mm^2$	Not specified	1280×960

Table 2.1. *Comparison of size, dynamic range and resolution of several different event-based sensors*

compare these characteristics of different event sensors for which specifications are published. The comparison of sensor size, dynamic range and resolution is presented in Table 2.1. As Table 2.1 shows, the sizes of the sensors are already small enough to be considered for mobile phone usage, and have been of such size for long time. Sensor resolution, on the other hand, has been steadily increasing over the years. It is hard to determine which resolution would be considered as sufficient enough to justify the event sensor mobile phone integration, but if the trend continues, we can expect to see sensors with even higher resolution in the coming years.

2.3 Use cases

Most of the event camera use cases are related to computer vision. Samsung's SmartThing Vision was the only consumer product containing DVS sensors that have been sold to customers, but the product was discontinued later, and it is not anymore available for purchasing [15]. The product is a home security device which can detect intruders or events where a person might injure themselves by for example falling in front of the sensors view, alerting other people connected to the smart home system. The decision to put event sensor into the product was advertised by arguing that it increased the level of privacy compared to RGB sensors. Security implementations of event sensors are further explored by [16], focusing on object detection in dark outdoor conditions. Detecting persons from the event data is especially useful, if that can be done faster than with other sensors. Autonomous automobiles is one example of an area where time is of essence when detecting objects, when moving at speeds which would result in serious consequences in the case of collision. Sokolova and Konushin have shown that gait detection is possible using event sensors and the accuracy is at par with the state-of-the-art RGB-based methods [17], and the results suggest that event sensors could be used in both pedestrian detection in automotive applications and security applications where person is identified by modeling the unique attributes of their gait.

Sarmadi et al. [18] have demonstrated that using event camera as data source, it is possible to reliably detect fiducial markers, which could be used for example as spatial references to inform autonomous vehicles about their position or directions. Examples of such cases are robots on the factory floor and unmanned aerial vehicles (UAV's). These

types of markers can also be detected via other types of sensors, but significantly less computing power is required with event-based solution, which is also able to detect the markers from higher velocity due to the asynchronous nature of the camera. The case of UAV's equipped with event cameras has been explored by Falanga et al. [19], and it was concluded that by reducing the latency of object detection to 3.5 ms it was possible to perform effective obstacle avoidance in speeds up to 10 meters per second.

In 2021, the United States National Aeronautics and Space Administration (NASA) conducted the first ever experiment of autonomous flying vehicle on extraterrestrial celestial body, when their Ingenuity helicopter flew independently on Mars [20]. If these types of experiments are continued in the future, using event camera for localization and mapping for those types of vehicles could be useful, since in those types of scenarios power is very scarcely available, and hence the lower power consumption of event camera compared to other sensors could provide some advantage. In the comparison between RGB and event sensors, the difference of power consumption is dependent on the amount of movement that the event sensor detects, since more movement leads to more events. When comparing event sensors to depth-sensing sensors that can be used in autonomous vehicles, using event sensors could result in up to 90 % less power consumption [21].

2.3.1 Mobile imaging

In the later half of 2010's and since, there has been steady increase in the amount of different camera sensors in mobile devices [22]. Data from multiple high-quality imaging sensors combined with more efficient neural networks and other such algorithms has made it possible to increase the capability of mobile devices to become viable alternatives to digital single-lens reflex (DSLR) cameras, which have seen decline in sales in countries such as Norway and Germany [23] [24] at the same time when mobile phone sales have soared.

Examples of sensors that provide additional functionality to the mobile imaging space are TOF sensors which can be used to estimate the distance between the sensor and the subject, and structured light sensor, which is perhaps most well known from Microsoft's Kinect sensor, which can detect movement and gestures by a person in 3D space. TOF and structured light sensors have already been included in devices by multiple different manufacturers, such as Huawei, Samsung and Xiaomi.

There are several applications on what event cameras could be useful on if they were included in mobile devices. For example, event data captured alongside other data could allow the generation of slow-motion videos after the capturing, in the post-processing phase. Rebecq et al. have demonstrated that event data can be utilized for increasing the attribute *frames per second* (FPS) for videos significantly, making even >5000 FPS possible [25]. Higher FPS in traditional slow-motion videos means more frames that need

to be saved, which leads to greater usage of storage space when capturing slow-motion videos with traditional methods. Due to the lightweight nature of the event data, it could be a viable option for slow-motion video creation in the future.

Other mobile applications could include motion- and handshake-deblurring, of which the latter is addressed in more detail in further parts of this thesis. Video and picture quality enhancement on content captured on dark conditions are also cases where event cameras could be used. Event cameras have also been proven to be efficient in iris-tracking by Ryan et al. [26], the motivation behind their study being to examine the possibility to utilize event data to monitor the ability of the driver when they are driving a car. However, the results also suggests the possible benefits of implementing event capturing capacity to the front-facing cameras on mobile devices to be used in applications where the user could interact with the device by solely blinking or moving their eyes. Event-based eye tracking can also be used for face detection, as shown by [27], although face detection and user identification based on facial features is already possible in dark conditions by methods implemented in some mobile phones that are being sold today. Gesture recognition in more general form by utilizing event data has already demonstrated as possible by Chen et al. [28]. Along with security, automotive and other such usages where the data could be useful, this type of detection could be used in mobile phones.

2.4 Event Data

A single data point in event data contains three different components; time (t), place (x, y) and the sign of light intensity change, also often called polarity (p). This format is called the *address-event representation* (AER) [8]. From a set of data points in AER format, several visual representations can be derived. One of the most straightforward ones is grayscale event frames, where a set of data points are plotted on an image which has the same resolution as the event-based sensor. It appears that most single-frame representations use either intensity representation defined by [29] or binary representation. In addition to the two, another type of *gradual representation* is introduced here. To emphasize the asynchronous nature of the event data, it is also possible to visualize the event stream in a time-continuous way. This representation adds time as additional dimension to the frame-based representation that was introduced previously, and is visualized in Figure 2.1 by [30].

In binary representation, for each pixel the sum of the event polarities on a given time window is calculated, and if the sum is negative, the pixel is portrayed as negative event. Positive sum is similarly portrayed as positive event. Pixels portrayed as negative events are plotted as minimum value (0) on the grayscale pixel value spectrum and pixels portrayed as positive events as maximum value (255). Neutral pixels where the sum is zero take the middle range value of 128. The method used for calculating the pixel value in

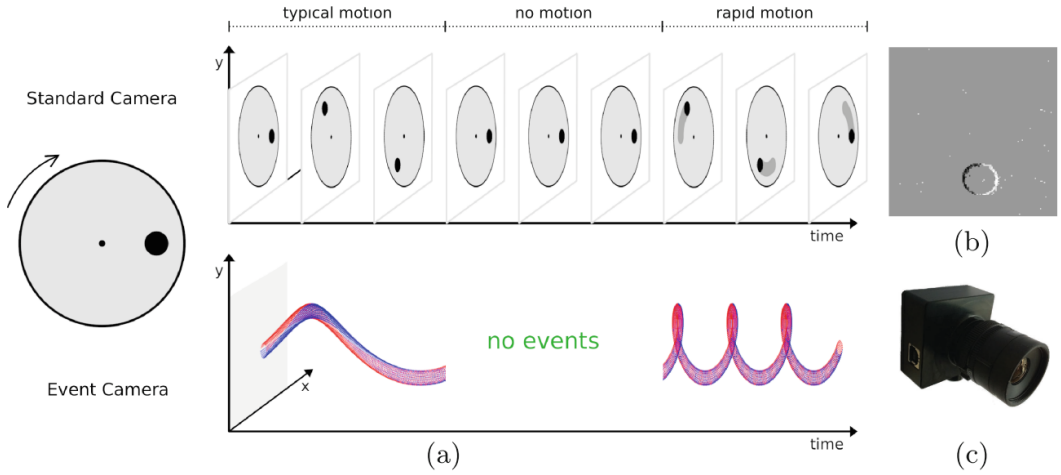

Figure 2.1. *Event stream representation including time dimension and distinction between outputs of traditional and event-based sensors, included in [30]. (a) depicts the event sream, (b) is the frame-based representation of the event data, a snapshot from the stream and (c) is a DVS event camera.*

this manner is presented in Equation 2.1.

$$E_{xy} = \begin{cases} 255, & \text{if } P \geq 1 \\ 0, & \text{if } P \leq -1 \\ 128 & \text{otherwise} \end{cases}, \quad P = \sum_{i=1}^{L} p_i \quad \text{and } p_i = \{-1, 1\} \tag{2.1}$$

In Equation 2.1, E_{xy} is the value of a single pixel in the event frame, L is the number of events for a pixel where the coordinates are (x, y), and p_i is the polarity of a single event. Gradual representation portrays events as values anywhere from 0 to 255, depending on what the sum of the polarities is during the given window of time and what is the maximum difference compared to the baseline value 128. The method for this calculation is visible in Equation 2.2.

$$E_{xy} = 128 + P, \quad P(I, x, y) = \frac{128(I_{xy} - 128)}{max(|I_{max} - 128|, |I_{min} - 128|) - 128} \tag{2.2}$$

In Equation 2.2, I stands for the image frame where all of the polarities for each pixel have been summed together, and where values can exceed 255 or be less than 0. I_{xy} is the initial value of the pixel in the coordinates (x, y) in I and I_{max} and I_{min} are the minimum and maximum values inside I. This representation contains more information about the scene, since it makes it possible to directly observe how many times the illumination of a given pixel has changed enough to trigger and event, and thus it is the type of representation that is used in the upcoming visual representations of event data and the experiments that are conducted. In some cases, negative events are portrayed as

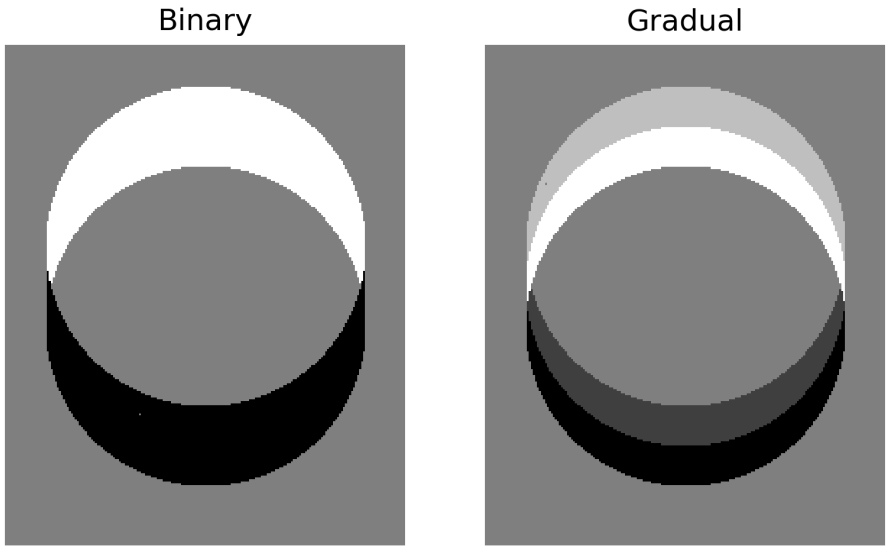

Figure 2.2. *Gradual and binary types of frame-based representation of event data.*

$p = 0$ whereas some cases $p = -1$ is used. In order for these representations to work properly, the latter notation is used. Visualization of binary and gradual representations is displayed in Figure 2.2.

As it was stated earlier in this chapter, biological systems reduce the information that is passed forward by significant amounts to avoid information overload in the brain. Similarly, the difference of the amount that is contained in raw RGB images can be compared to the amount stored in event data. If the resolution of a raw RGB image is $w \times h$, the amount of information can be expressed by Equation 2.3.

$$bits = w \times h \times b \times 3 \tag{2.3}$$

Where b is the number of bits per pixel on each of the three channels (red, green and blue), for example 8, 16 or 32, with larger number of bits per channel resulting in more realistic colors. As for the event data, the number of bits stored in a single event is expressed in Equation 2.4.

$$bits = b_w + b_h + 32 + 1 \tag{2.4}$$

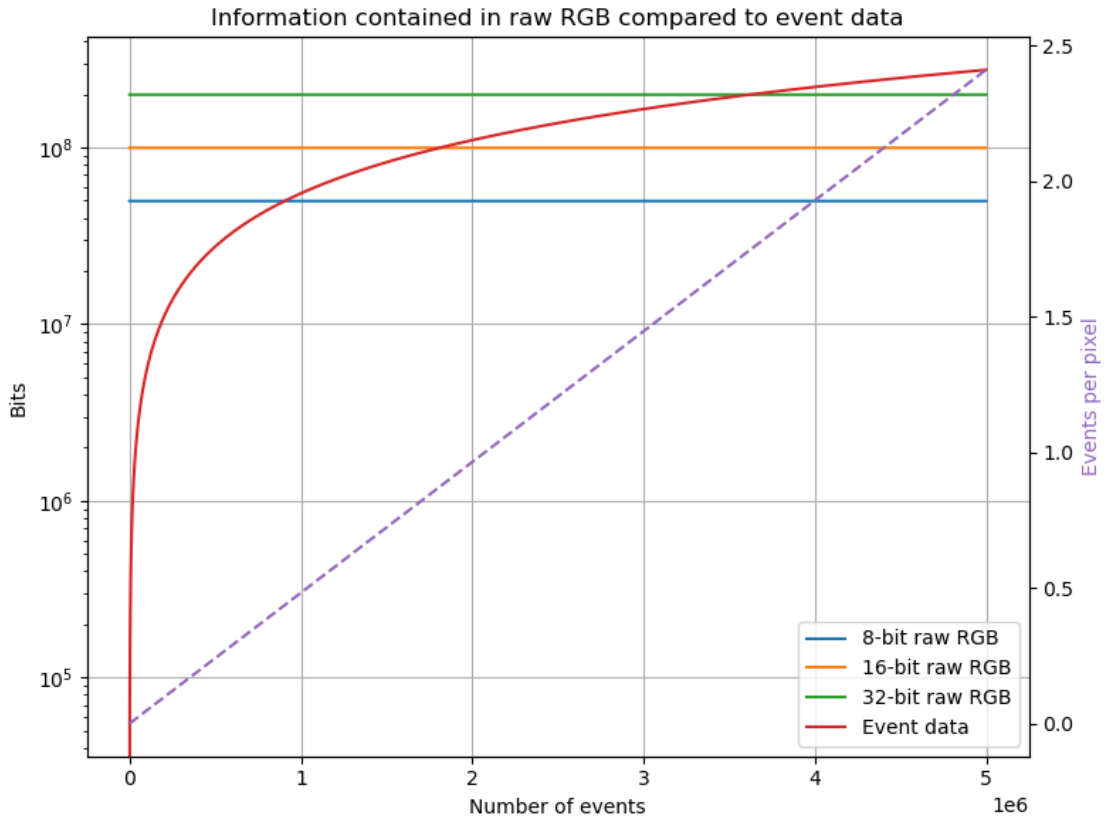

Figure 2.3. *Amount of information contained in raw RGB and event data where the number of events per pixel varies, both from sensor of size* 1920×1080*. Number of bits is in logarithmic scale.*

Where b_w and b_h are the amounts of bits that are required to store the information about the event width and height, respectively, and which is depend on the sensor size. Timestamp of an event is considered to be stored at microsecond precision, resulting in 32 bits, and polarity of an event requires a single bit. Using Equations 2.3 and 2.4, a comparison between information contained in event and raw RGB data is conducted. Assuming a sensor size of 1920×1080 for both event and RGB sensors, the comparison is visualized in Figure 2.3.

As the Figure 2.3 shows, there has to be almost two events per each pixel in the event data for the amount of bits required to store it to match the amount of bit required in storing a 8-bit raw RGB of the same resolution. However, as the event sensors have already been proven to provide data with such precision that many different applications have been made possible even without reaching the full high-definition resolution, and taking into consideration that most of the RGB sensors already have a resolution that is higher than full high-definition, it is reasonable to suggest that in reality the sensors in a setup where both event and RGB sensors are used will not have same resolution, and therefore the difference in the amount of bits required to store the information is even higher.

2.4.1 Simulators

Since event cameras are still quite uncommon, and not too many event data data sets are freely distributed online, it follows that all the people who would like to participate in investigating and developing applications and algorithms that utilize event data may not be able to do so. For this reason, open source event data simulators such as ESIM [31] and AirSim [32] have been made publicly available. Luckily, from the technology development point of view, the amount of event datasets have been increasing in the last years, and now there exists several freely distributed datasets for different areas, such as automotive [33] [34] and UAV's [35]. However, event simulators make it possible to generate event datasets from any RGB video or image sequence, increasing the scope of available data significantly.

The working principle of ESIM is as follows; series of consecutive RGB image frames are taken, and temporal upsampling is applied to interpolate new frames between the original images by arbitrary temporal resolution. This way the illumination signal in high temporal resolution can be approximated to mitigate the output of a real event sensor. Then, consecutive frames from the new set which contains more frames are compared to each other, measuring the change of illumination in each pixel. If the change exceeds the threshold parameter being used, an event is triggered and stored to the output.

Another benefit of simulated event data compared to data acquired by using a system consisting of a real event sensor alongside the RGB sensor is that in the event data created by the simulator is perfectly aligned with the RGB data that is used as an input for the simulator. In cases where the data is acquired by two spatially separated sensors, registration is required. By registration, misalignment caused by the distance between the sensors and different field-of-view (FOV) of the cameras is compensated so that the two images captured by two different sensors become aligned. In addition to the registration, using real sensors would require stereo camera calibration to diminish the effect of different lens distortions in the two camera modules.

3. DEBLURRING EXPERIMENTS

Handshake blur is the type of blur that appears in the image when the imaging sensor is moving during the exposure. It is important to note the difference between handshake and motion blur, of which latter is the type of blur which happens when the subject in front of the sensor is moving during the exposure time, rather than the sensor itself. The blur types can also occur simultaneously, when both the sensor and the subject are moving during exposure. Several different solutions have been developed to combat this problem, including the usage of optical image stabilization (OIS), and gyroscopes that track the movements of the camera and then use that information for reconstructing the trajectory that the device moved during exposure, and then use this information in postprocessing to compensate for the relative movement between the sensor and the scene [36]. Also, purely software-based methods such as [37] which do not rely on any external hardware or source of data have been developed to combat the issue.

As mentioned in previous chapter, event cameras have several potential use cases. Deblurring of motion blur using event data has been demonstrated as possible task by [29] and [38]. It would therefore be reasonable if the event data could also be utilized for reducing the blur caused by imaging sensor motion as well. Here a study is conducted on examining the possibility to utilize an existing neural network based method which was initially proposed for simultaneous deblurring and denoising, and to replace one of its two RGB inputs with event data. Since event cameras gained their initial inspiration from the way biological retina passes information forward and neural networks mimics the processes in which biological neurons take part, these types of applications can be considered to be a part of an interdisciplinary field called *neuromorphic engineering*, which studies the utilization of bio-inspired models to solve different engineering problems.

LSD2 network developed by Mustaniemi et al. [39] is based on U-Net, which is a neural network architecture used for simultaneous deblurring and denoising. U-Net was originally developed by Fischer and Brox to be used in the segmentation of biomedical images [40], and the name comes from the fact that the network architecture contains a contracting path and expansive path, of which the latter takes a curved route in the network architecture visualization, giving it it's U-shaped form. As mentioned, original LSD2 can perform both denoising and deblurring. It takes two inputs, one taken with short exposure time and which contains noise but is sharp, and other which is taken with long

Figure 3.1. *Sample images from the MIRFLICKR dataset [41].*

exposure time, is blurry but does not contain noise. Because the noisy input image is replaced with event data in the following experiments, denoising part is irrelevant for the task and focus will be on the deblurring. LSD2 model is chosen due to the analogous nature between the data it takes as input and event data. Both event frames and the short-exposed RGB image that the original LSD2 network takes as input are snapshots of the scene with low temporal resolution. The low temporal resolution of those two allow them to contain the details which are missing in the blur, and are thus considered helpful for deblurring purposes.

MIRFLICKR dataset [41] is chosen as the source of RGB images. The dataset is a collection of 100 000 images with a wide range of different types of content from the online photo hosting service Flickr. The dataset also includes a set of one million photos, but for the case of a image quality enhancement neural network training, 100 000 photos is considered as a size that is sufficient enough due to the high variance among the contents that the dataset contains. Randomly chosen sample images from the MIRFLICKR dataset are displayed in Figure 3.1. As it can be seen from Figure 3.1, the images are high quality and the variance among contents is large, which are considered as benefits from the perspective of training a model.

Some modifications are required for the LSD2 model to work with event data. Size of the input layer of the model was changed to be adjustable for different sizes and shapes of inputs. Generators which feed training data to the networks also needed to be created to support event data formats.

Parameter	Explanation	Value
Trajectory size	Maximum square size of the trajectory area in pixels	30
Anxiety	Amount of shake, on scale from 0 to 1	0.25
NumT	Number of points along trajectory	10
Max total length	Maximum sum of euclidean pixel distance between all the points on the trajectory	20

Table 3.1. *Random motion trajectory parameters*

3.1 Data generation

Training data is needed in order to train the neural network to model the correspondences between the input event and RGB pairs. It consists of three parts; the ground truth, or the ideal output, which is the original sharp RGB image, the blurry RGB image and the event data. Blurry RGB image and event data act as the inputs from which the network should be able to compute an image that resembles the original sharp image. In order to generate the handshake blur effect for the blurry input images, the motion of a moving imaging sensor should be simulated. This is done via process called random motion trajectory generation. The method used here is based on the implementations by [42], which is considered suitable due to the fact that the methods for trajectory generation were developed for artificial blur purposes in the first place. Some parameters are required to be set for the trajectory generation, including its size and length. The parameters and the values used are presented in Table 3.1.

After the points of random motion trajectory are generated, they will be applied to the images that were obtained from the MIRFLICKR dataset. The trajectories are applied to the images by setting the first point of the trajectory at the center point of the image. Images from MIRFLICKR dataset vary in size, which means that in order to standardize the input size, the trajectory will act as a path for a moving window from which samples of constant size will be obtained by using a MATLAB script that conducts camera movement emulation [43]. A window of 256x256 pixels is chosen for practical reasons. The window size is big enough so that in most cases, there remains objects inside the window that are suitable for deblurring. The small size also makes the process of generating data and training the network used for deblurring faster. The first image from the center of the image is chosen as the ground truth, which will be used in the training to teach the network what the output should look like. Windows moving along the trajectory on top of an image is visualized in Figure 3.2

Some original images from the MIRFLICKR dataset are discarded during the data generation due to their small size. Discarding happens only in cases where some part of the window that is being moved in order to get the blurry image moves beyond the borders of

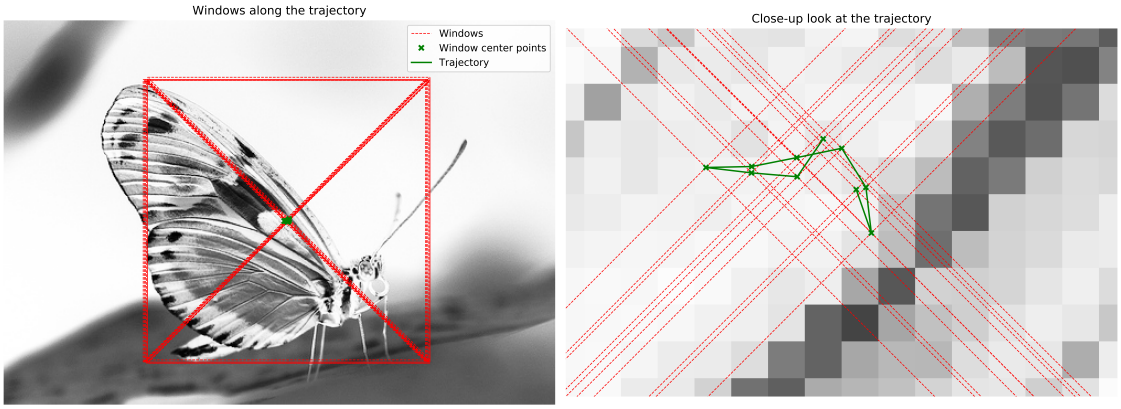

Figure 3.2. *A 256x256 window moving along the random motion trajectory on top of an image. Image is presented in black and white for visualization purposes.*

the original image at some point of the trajectory. This leaves a total of 88 293 remaining images from the original 100 000, which is considered a sufficient amount for the task.

Additional frames could be added inbetween the consecutive original frames via upsampling in order to build images with more smooth blurry areas which resemble a more realistic scenario. This step is not conducted here because there are relatively few images with non-smooth blur containing the types of edges that become apparent when objects in the images that are being averaged are too far apart from each other. The same problem can be tackled by reducing the spread of the random motion trajectory by adjusting the parameters presented in Table 3.1 so that the points are closer to each other, but still far enough from each other for the averaged image to contain blur.

The ten images obtained along the trajectory from slightly different positions in the original image are used as input for the ESIM, which generates event data for each transition from one image to the next. Event data from the simulator is given as AER format as explained in Chapter 2, containing timestamp, event location and the event polarity. This data is divided into chunks by splitting events according to their timestamp, and those chunks are being used for event frame generation. Different amounts of event frames per each transition should be tried, to see which input format is optimal for training to achieve best results. First event frame from the simulator output has quite few events, and it is discarded. Now we have nine transitions from one image to other, which means that number of event frames in the stack should be divisible by 9 for the amount of event images per transition to be equal. Smaller sizes of event stacks could also be used, but if the time interval used for constructing each event frame would be too large then the sharpness of the details in the event images would decrease. Two event data stack sizes, 9 and 18, are chosen to be used in the trials.

Event simulator requires thresholds for positive and negative events as input, as explained

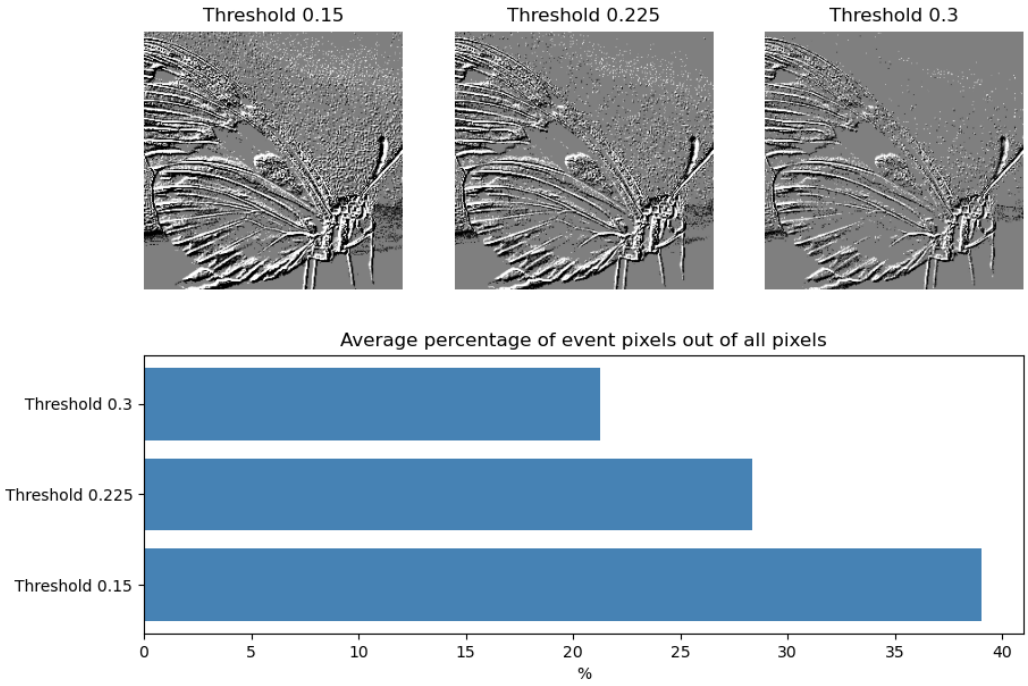

Figure 3.3. *Different event thresholds visualized.*

in Chapter 2. The thresholds represent the so-called sensitivity of the sensor, meaning that with lower threshold, smaller change of illumination of the light perceived by the event sensor is required in order to trigger an event. Lowering the threshold comes with a trade-off; more details can be seen, but more noise appears as well. This is illustrated in Figure 3.3. It can also be seen that halving the simulator threshold almost doubles the percentage of event pixels in the image. There can be different thresholds for positive and negative events, but here those two are kept equal in all instances where the simulator is utilized for event generation.

After the data generation is complete, 88 293 samples are obtained with each sample containing a stack of event frames, one blurry image and the original sharp image. Resulting data and the process for generating it is visualized in Figure 3.4.

The input provided for the network consists of three color channels (red, green and blue) from the blurry RGB image stacked on top of all the event frames that have one channel each since they are in grayscale format. For the case where motion trajectory contains n_1 points and n_2 event frames are constructed per transition, the size of the input layer will be $256 \times 256 \times (3 + n_1 + n_2)$.

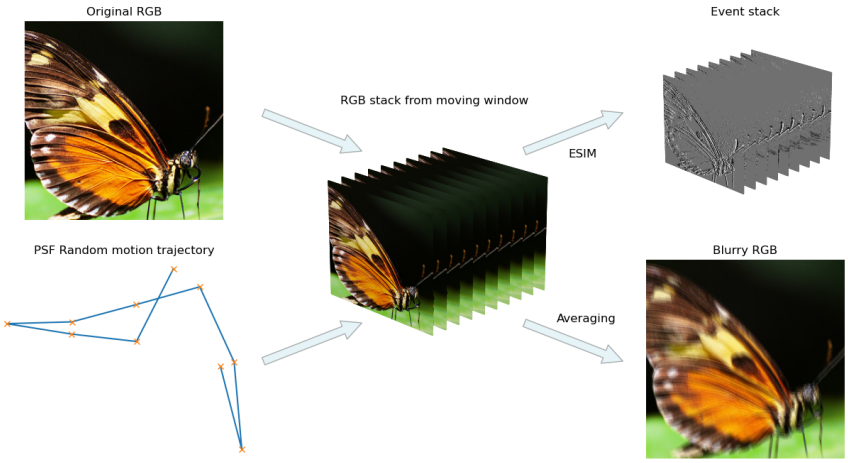

Figure 3.4. *Process of creating handshake motion, event stack and blurry image from single RGB image.*

3.2 Training

Several different combinations of input data and training parameters should be tried in order to find how event data should be formatted in order to yield best results. Evaluation of the outputs after training is done by *peak signal to noise ratio* (PSNR) and *structural similarity* (SSIM) [44], which are widely used metrics for image quality assessment. During training, the evaluation of model performance is done by using the loss function of the model, which in this case is mean squared error (MSE). The way the MSE, PSNR and SSIM are calculated are presented in Equations 3.1, 3.2 and 3.3, respectively.

$$MSE(image_1, image_2) = \frac{1}{XY}\sum_{i=1}^{X}\sum_{j=1}^{Y}(image_{1_{ij}} - image_{2_{ij}})^2 \tag{3.1}$$

$$PSNR(image_1, image_2) = 10log_{10}\left(\frac{255^2}{MSE(image_1, image_2)}\right) \tag{3.2}$$

$$SSIM(image_1, image_2) = \frac{(2\mu_1\mu_2 + C_1) + (2\sigma_{12} + C_2)}{(\mu_1^2 + \mu_2^2 + C_1) + (\sigma_1^2 + \sigma_2^2 + C_2)} \tag{3.3}$$

In Equations 3.1, 3.2 and 3.3, X and Y are the image dimensions, in this case both equal to 256. μ is the average image value, σ is the image variance, and C is a parameter defined in Equation 3.4.

$$C_1 = (k_1 L_1)^2, \quad C_2 = (k_2 L_2)^2 \tag{3.4}$$

Experiment ID	Event threshold	Event stack size
E01	0.300	9
E02	0.300	18
E03	0.150	9
E04	0.150	18
E05	0.225	9
E06	0.225	18

Table 3.2. *Event data formats on initial experiments*

In Equation 3.4 k_1 and k_2 are constants, 0.01 and 0.03, respectively, and L is the dynamic range within the image, defined by being the difference between maximum and minimum values. From Equations 3.2 and 3.3, it becomes evident that for both PSNR and SSIM, larger value corresponds to a higher similarity between the two images that are being compared. Since the images here contain three channels; red, green and blue, the value is calculated as the average of the values for MSE, PSNR and SSIM are calculated using all three channels. Average PSNR and SSIM values are also calculated using the blurry images from validation data and the sharp ground truth image, to give an impression on where the baseline is when examining the development of the models in training. The validation data baselines are 18.20 for PSNR and 0.44 for SSIM.

It is worth noting that SSIM has also received criticism on its accuracy, especially on the case of evaluating RGB images. Nilsson and Akenine-Möller have explained that image quality that is perceived by humans can vary considerably of that which is mathematically calculated via SSIM, making the metric somewhat unreliable [45]. Nevertheless, it is still among the most used and efficient metrics available, and hence it will be used here. Due to concerns of unreliability, subjective evaluation is needed together with the objective metrics.

3.3 Experiment I: Event parameters

Training is done in two phases in order to find what parameters are the most important for successful deblurring. In the first phase, different event thresholds and event stack sizes are tried. Experiments and their corresponding labels are presented in Table 3.2. After the initial trials, it is observed which parameters result in best results both objectively and subjectively, and after that the training is conducted again while keeping those parameters which were found out to work best as constants, and by tuning other aspects of the training.

Each experiment in the first phase is trained for 100 epochs with batch size of 1000 images. Learning rate does not vary between experiments, as it starts at 0.00005, and

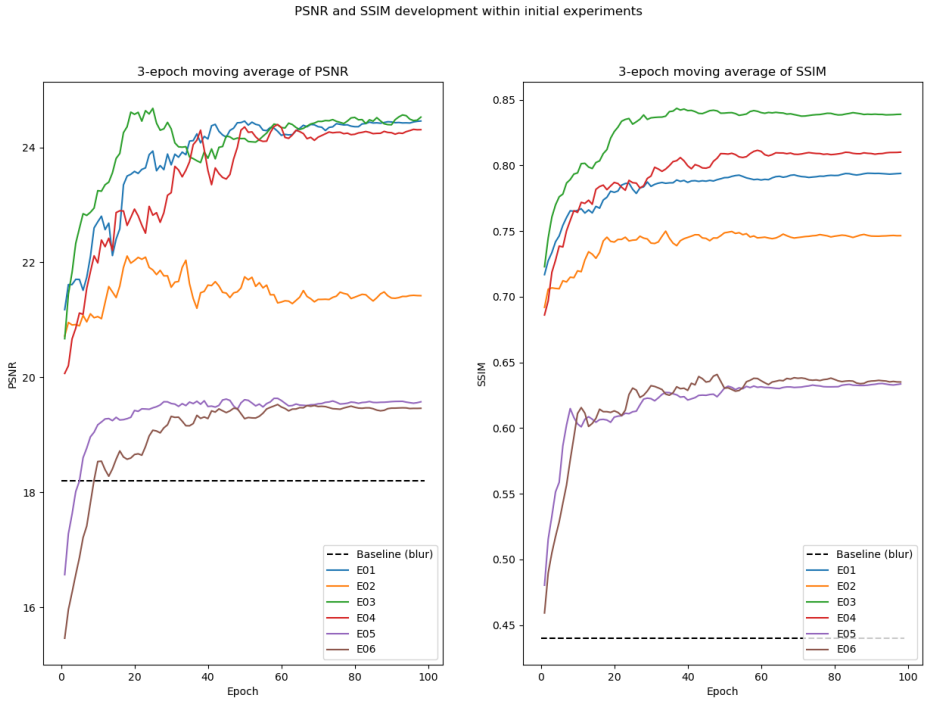

Figure 3.5. *Image quality on validation data during the training of initial experiments, as measured by PSNR and SSIM.*

halves every 10 epochs. From these initial experiments, best parameters for event threshold and event stack size are chosen for further experiments with other network parameters. Average PSNR and SSIM values calculated after each epoch from a set of outputs generated using images from validation data is presented in Figure 3.5. The set of validation data images remains constant between epochs and experiments, so the numbers are comparable. In Figure 3.5 and all following figures of the same format, three-sample moving average is used to smooth the curves for clearer interpretation.

Learning curves by PSNR and SSIM metrics show that the models with 0.15 threshold perform best with this type of data, and hence it seems that the benefit from increased accuracy in the event data is great enough to offset the possible downsides caused by increased noise, which was visualized in Figure 3.3. The curves also imply that the effect of the threshold to the model performance is not unambigious, since the threshold of 0.225 results in worse performance than both 0.30 and 0.15. The curves of PSNR and SSIM do not provide much insight into the actual image quality of the outputs from the human perspective, and hence visualization is needed. Deblurring is performed for a set of images from the test data, and some samples from the test data outputs are displayed in Figure 3.6.

The results presented in Figure 3.6 give a promising picture of the possibilities of utilizing event data for deblurring. The blur is clearly reduced in the outputs, but otherwise the

Figure 3.6. *De-blurred images from testing data using models developed in initial experiments, and choosing the ones with highest PSNR and SSIM values.*

results are poor quality. The dynamic range is lower than in the original image, and some artificial noise can be seen in all of the outputs. The top-right image in the figure shows that the blur region is still present in the output, and it can be seen especially well since there is high contrast. For further experiments, 0.15 threshold is kept since it performed best by both PSNR and SSIM metrics and subjective evaluation.

3.4 Experiment II: Network parameters

For further experiments, threshold 0.15 is kept constant like previously mentioned. In the experiments, variables are the initial learning rate, learning rate decay and batch size, although different batch size is tried only once. For detailed descriptions of experiment

Experiment ID	Learning rate at start	Learning rate decay	Notes
E07	0.0005	0.75 * every 10th epoch	
E08	0.00005	0.50 * every 10th epoch	
E09	0.0005	0.75 * every 10th epoch	
E10	0.0005	0.95 * every 10th epoch	Try out significantly smaller batch size (100); bad results
E11	0.0005	0.95 * every 10th epoch	Subjectively best results
E12	0.005	0.75 * every 10th epoch	Try out significantly higher learning rate; bad results
E13	0.0001	0.75 * every 10th epoch	

Table 3.3. *Training parameters on further experiments*

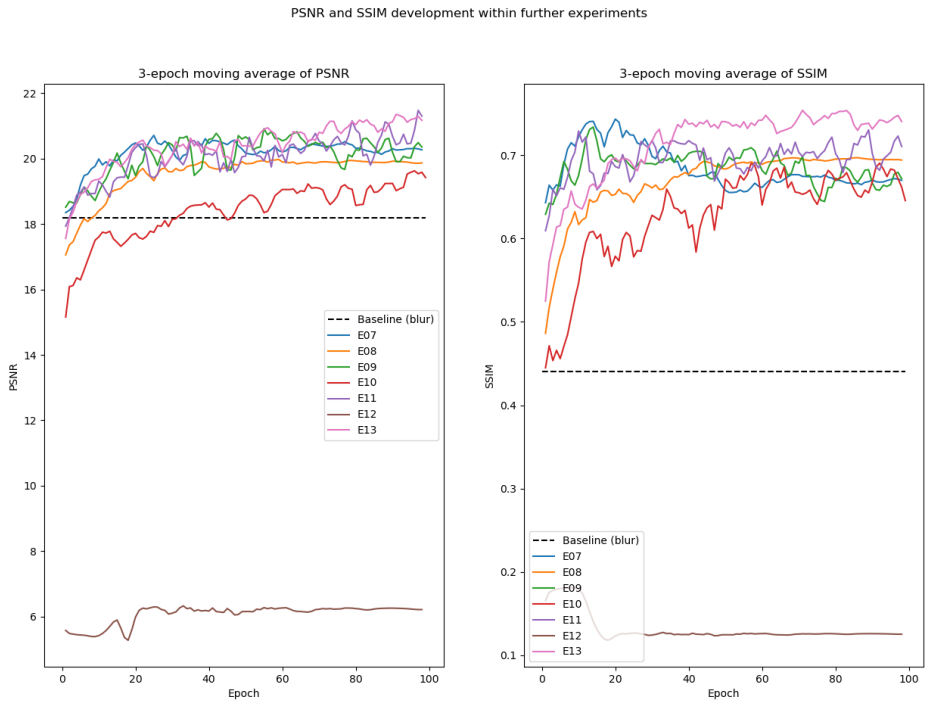

Figure 3.7. *Image quality on validation data during the training of second phase of experiments, as measured by PSNR and SSIM.*

parameters, see Table 3.3.

The training progress is visualized in the same format as previously. Figure 3.7 shows that the difference between model performance is smaller than in the initial stage, as expected, when not including one outlier (E12). Interesting observation can be done when comparing figures 3.5 and 3.7; the values for PSNR and SSIM are lower in the latter experiments, and thus can be considered worse than the results obtained in preliminary experiments. However, the results presented in Figure 3.8 show that the outputs are less noisy, and have greater dynamic range than their counterparts generated by models trained in the initial phase. In addition to that, the PSNR and SSIM values calculated within the testing data images are better with the models developed in latter experiments, even though the same metrics were better for the initial models in the training phase. The output images portrayed in Figure 3.8 are picked by the highest PSNR and SSIM values, when comparing the output images of all experiments detailed in Tables 3.2 and 3.3.

Figure 3.8. *De-blurred images from testing data, picking best by PSNR and SSIM among all models developed.*

Even though the image quality on testing data increased within second phase experiments, shortcomings of same type can be seen in the images in Figure 3.8 as well, but as having a minor effect compared to images in Figure 3.6. Resulting test image outputs are examined for all the models that were trained, to see if subjectively perceived quality matches the objective metrics. For both sets of testing images, image quality that is subjectively perceived by the author among the outputs of all models is aligned with the PSNR and SSIM based evaluation, except in the case of the top row in Figure 3.8, where the model E11 looks subjectively better than the ones picked by highest PSNR and SSIM. Detailed visualization of model E11 performance on an image from testing data is visualized in Figure 3.9.

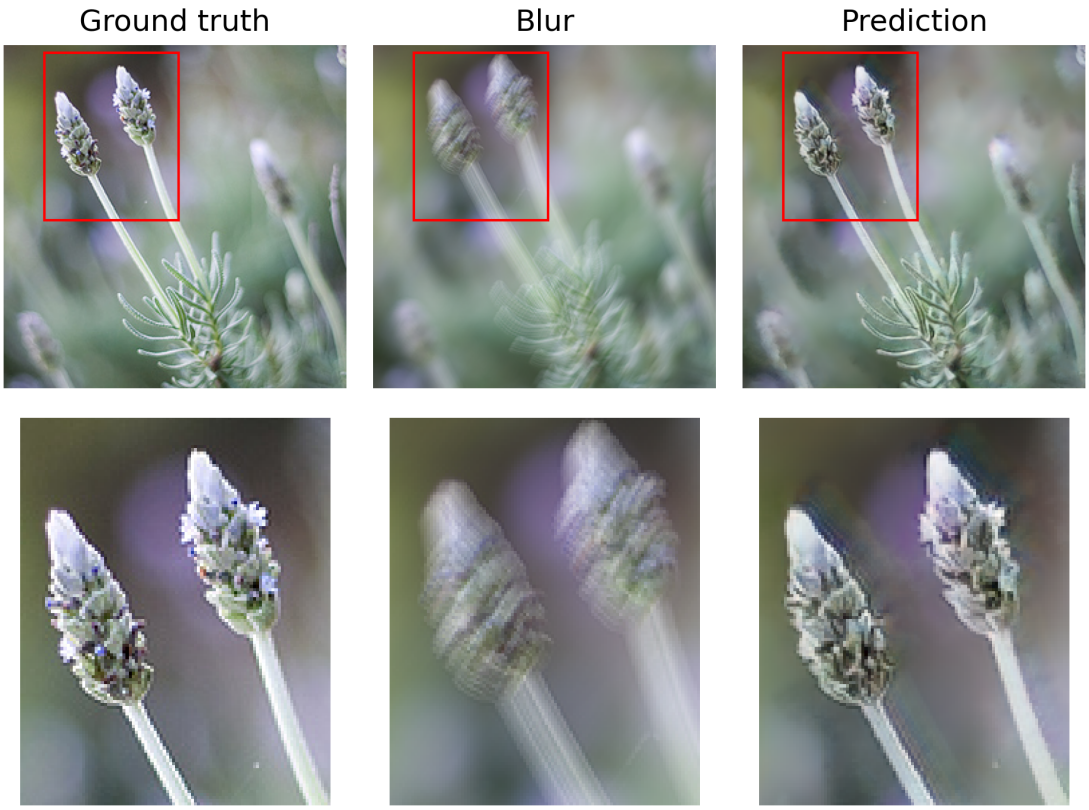

Figure 3.9. *Detailed visualization of model E11 performance on an image from testing data.*

As it can be seen from Figure 3.9, although the predicted image is a lot sharper than the blurry input, some defects can be observed. Deblurred image has lower dynamic range, and the area that was covered by the blur in the blurry image is still mildly visible in the output.

Evaluation with external data should be additionally conducted to assess how well the deblurring works with other type of data than the type for which the generation was visualized in Figure 3.4. It will be especially useful to see whether it is possible to use a blurry image for which the sequence of images are not generated by the random motion trajectory, but rather by natural movement of the camera.

For external evaluation, GoPro dataset by Nah et al. is used, since it has been proven to be useful data set for deblurring applications [46]. From that data set, a single video is chosen for evaluation, and blurry image is generated by averaging ten consecutive frames from a video in the dataset which is shot at 200 FPS. ESIM is used again to generate event data for that image sequence. The video chosen for evaluation contains a license plate of a car, which is an useful case since it can be easily seen if the text becomes readable after the deblurring. Event data is generated in several different forms so that for each model, event data is created with the same stack size and event threshold that the model

Figure 3.10. *Deblurring performance on GoPro data by [46].*

was trained with. The outputs that achieved best PSNR and SSIM are presented in figure 3.10, alongside the output with subjectively best quality. There are several important notions that can be taken when examining the Figure 3.10. For one, the blur in the blurry image is worse than it was within the blurry images generated from the MIRFLICKR data. Despite that, the text becomes readable even though the image quality is otherwise poor. Second important notion is that in this case, the model achieving best PSNR and SSIM is clearly performing worse than the one that has been picked as subjectively best by the author.

3.5 Summary

From the model output results, it can be seen that event data can be utilized in this setting for handshake deblurring. There are few noticeable shortcomings on the model performance, perhaps the most obvious of those being that the colors of the outputs are less saturated and lack the dynamic range that the ground truth images have. This might be caused by the grayscale format of the event images, a characteristic that the model perhaps omits and passes forward to the output. As it was shown in Figure 3.10, developed deblurring method can also be applied to a dataset where the blur images are not generated by random motion trajectory as seen in Figure 3.4, but rather from authentic and natural movements of the camera, although this was demonstrated with only a single sample. Same shortcomings are present in the results of deblurring of this type, including lower dynamic range. Dynamic range issue is visualized in Figure 3.11, where the test data predictions from model that was considered best performer (E11) is compared to the

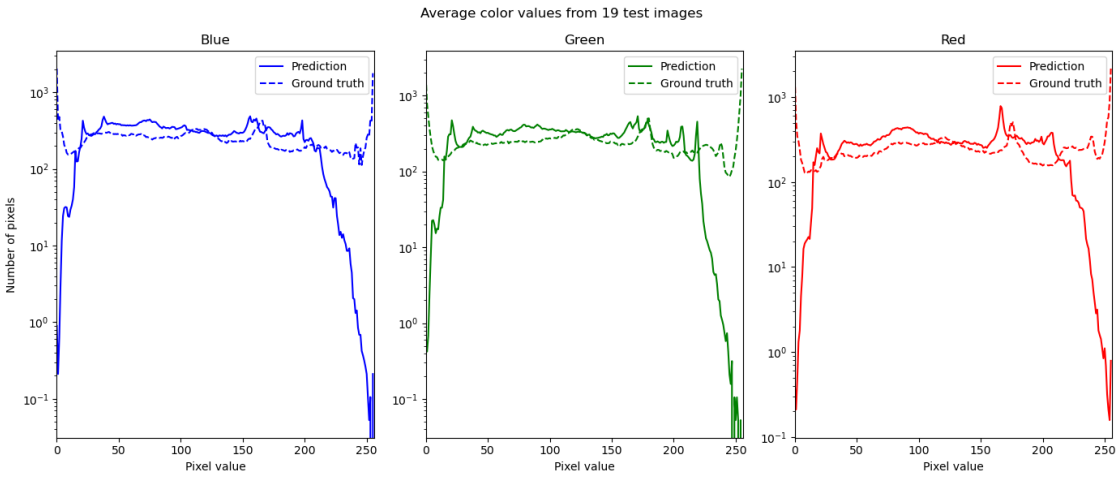

Figure 3.11. *Average color intensities of predicted and ground truth test data images with model E11.*

ground truths from the perspective of color value distribution among the images. It can be seen that the high and low end of the color spectrum have very low intensities among the output images, contrary to the ground truths, where the peaks are on both edges of the color intensity spectrum. The effect is stronger on the higher end of the color scales, meaning that the model performs better on dark than bright targets.

The dynamic range issue that is present in the results could be perhaps addressed in further research by integrating guided image filtering (GIF) step to the network architecture. Marnedires et al. [47] have developed a version of the U-net where GIF can be used for dynamic range expansion also known as inverse tone mapping (ITM), called GUNet. The GUNet architecture also tends to reduce artifacts in the output, but in the experiments conducted here, the artifacts are mainly remainders of the blur area that was removed when sharpening the image, which makes it unlikely that GUNet would be able to address this issue.

It is worth noting that these experiments were done only to demonstrate the possibility of using event data in LSD2-based application, and hence some things such as the perfect alignment between the RGB and event data were taken as granted. The sensitivity that is equal to the event threshold that was providing the best results might be impossible to achieve with event sensors available at the market today, without creating excessive noise. Although neural network based approaches are mostly what contemporary studies on image quality enhancing are focused on, event data could be also helpful with other methods, since it could be used to calculate the blur kernel of an image.

There exists almost endless possibilities on different parameters and adjustments that could be tried in order to drive the model performance closer to ideal. The trials conducted here and the results that were presented can offer some idea on which direction to move in order to achieve better performance. The smallest threshold of the three used was proven

to be the best, so decreasing the threshold even further down could yield improvements, even though it might cause the scenario to become too unrealistic, as it was previously mentioned. The slowest pace for learning rate decay was found to be the best among all models that were trained, and therefore in further research it should be considered to further decreasing of the learning rate decay. It should also be considered to train the network with data where the event threshold is varying among the images. This would resemble a more realistic scenario, since the ideal threshold is dependent on the scene, and thus different amounts of event data would be available for images captured in different scenes. Other possible modifications could include for example changing the ground truth image from the start of the artificial motion sequence of the images to the middle.

4. EVENT CAMERA TECHNOLOGY LIFE CYCLE ANALYSIS

The study of life cycles of from the industrial perspective is highly concentrated around product life cycles (PLCs), leaving the study in the field of technology life cycles (TLCs) in a significantly smaller role. When searching for scientific literature on technology life cycles on Google Scholar yields approximately 10 100 matches at the time of writing, similar search for literature about product life cycles (PLCs) yields approximately 237 000 matches. Similar but slightly smaller difference was discovered by Taylor and Taylor in 2011, using Abi Inform as the source [48]. This difference might be caused by the fact that a single product has narrower scope than the technology it is based upon, and therefore there is more variation in that space and more topics for research. While understanding both of these topics is important in order to achieve efficient and sustainable business practices and make informed decisions at the management level, here the focus will be on the technology life cycle and not on any individual product, event though the concepts of TLC and PLC are interlinked. From the management perspective, TLC analysis offers insights that can be helpful when making strategic long-term investment decisions in research and development (R&D) activities.

In this chapter, the technology life cycle models and different indicators and metrics are used to determine the current phase and future prospects of the event camera technology. Patent data will be used to gather information of event camera technology and few other technologies that can be used for comparison and validation of the models and methods that are being used for the analysis.

4.1 Technology life cycle model

The literature considering technology life cycles is not coherent in a way that there is not an established consensus on what models and terms to use, and although some models have been more widely adopted and used in research, no universally accepted model have yet been established, as pointed out by [48]. However, the S-curve has established itself as the dominant graphical representation of the technology evolution from the life cycle perspective, even though there is variance considering what exactly the S-curve portrays [48].

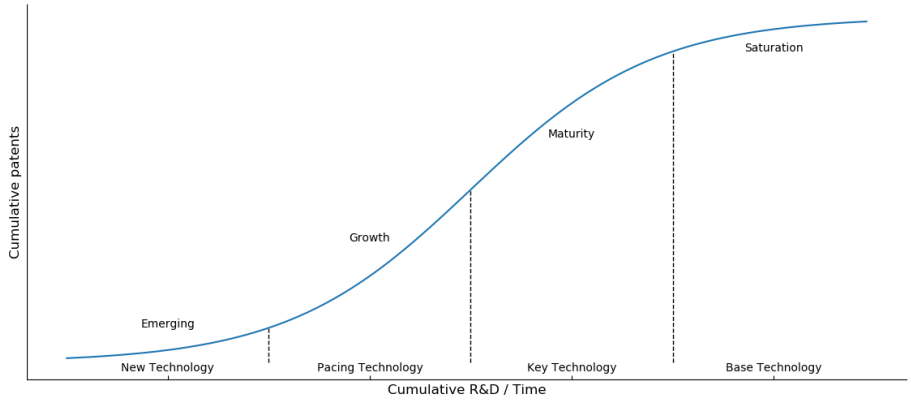

Figure 4.1. *S-Curve of technology life cycle, based on illustration by [49].*

S-curve representation of technology life cycle model, where the accumulated number of granted patents acts as the metric for inflection, was introduced by Ernst [49], and is displayed in Figure 4.1. The S-curve representation is used widely in modern research about TLC stage determination and forecasting, and it will be used here as well. Determining which metrics are adequately representative of performance of different types of technologies is difficult, as noted by [50]. Benefit of using accumulated patents compared to for example, accumulated amount of sales, is that patent data is available for observation earlier in time. This makes it more suitable for an analysis where the goal is to predict future trends, and the observations are ideally made as early as possible. Other advantages of using patent data in TLC analysis is that it is publicly and freely available, thus giving the analysis a high cost to benefit ratio. When using the logistic growth function that is depicted by S-curve, it is assumed that the variable under inspection starts from zero and has some upper limit which it reaches in some point in time. In TLC context and patents as the model variable, the thinking is that when the amount of total patents increases, the total knowledge behind the technology increases as well, making it possible to innovate even further, taking advantage of the established knowledge. This thinking is in line with the initial exponential growth, which is followed by stagnation when the full potential of the technology has been reached.

Several models have been proposed to represent evolutionary characteristics of technologies. Division of technology life cycle into four distinct stages has been widely adopted. In the early literature the TLC was interpreted as cyclical model which consists of four different eras; first era of ferment, second of emergence of dominant design, third of incremental change, and finally the fourth era of discontinuity of the technology, which is then again followed by the era of ferment [51]. In further research, S-shaped curve has become more commonly used, even though it lacks the cyclical visual representation that is present in the term life cycle. In S-curve models, eras have been redefined more simply

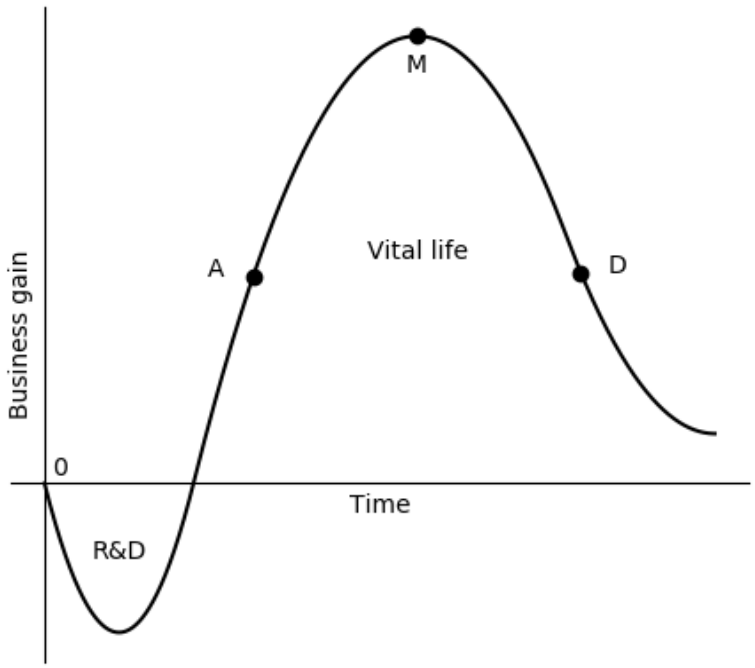

Figure 4.2. *Technology life cycle curve from business gain perspective, based on illustration by [54].*

as different stages, which are generally called emergence, growth, maturity and saturation. TLC stage naming convention is not something that is universally agreed upon, even though they represent similar characteristics in most of the cases. Sometimes stages get called different names than previously mentioned, for example the emergence stage can be called initiation [52] or embryonic stage [48], and saturation can be called decline [53]. Also, stages can be split, such as in [53] where growth stage has been divided into preliminary and real growth stages.

TLC from the business gain perspective is presented in Figure 4.2. At the early phases, losses are inevitable since resource requiring R&D activities are needed to be conducted in order to make further, commercially viable development possible. If the results during the R&D phase are good enough that the confidence among investors encourages further investment, eventually some products are developed and the investment costs are gradually covered. Point *A* on the graph stands for *ascent*, *M* for *maturity* and *D* for *decline* [54]. The four stages are analogous to the ones presented in Figure 4.1 along with the S-curve model.

4.2 Bibliometric analysis

Several sources for patent data are available. In this thesis, patent data is gathered from the United States Patent and Trademark Office (USPTO) database [55] for two reasons. First, searches from that database seem to provide more results per query than same searches from other databases such as Espacenet [56], which is the database owned and operated by European Patent Office (EPO). Second reason is related to the format that the patent documents are presented in at each source, and which is explored in the next section.

Using USPTO as the data source might bring up some undesirable skew in the data that is the subject of examination. Criscuolo has studied the home advantage effect which manifests in a way that domestic applicants are disproportionately represented in the domestic patent space compared to their foreign counterparts [57]. However, since the United States is the largest technology market in the world [58], most innovators worldwide desire to protect their intellectual property on that market, and that gives some explanation on why the disproportional majority of domestic applicants is smaller in the United States than in Europe, as observed by Criscuolo. Despite the effect, the analysis from Criscuolo comes to a conclusion that even with the domestic over-representation in the data, patent data from both both EPO and USPTO does offer a reliable picture into the international status of innovation in different technologies.

4.2.1 Patent data retrieval

Unfortunately, USPTO database does not offer an application programming interface (API) from which to access the patent documents in a programmatic manner, or a tool which would allow downloading all documents that match a given query at once. For this reason, programmatic browser-based method is needed in order to access the patents and download them. Python scripts were created to extract patent documents matching given queries from USPTO database in .html format recursively. Executing the searches in programmatic manner is possible using a Python package called Selenium WebDriver, which allows interacting with .html elements such as forms and buttons on a web page [59], and thus making it possible to acquire the search URL and the search results page containing links to patent documents matching the query by filling and submitting the query form, which is implemented via plain HTML. Finding the patent document links from the resulting web pages is done via BeautifulSoup Python package, which allows accessing the website elements systematically [60].

A distinction between the full patent documents and the documents retrieved should be made. The patent web pages which are downloaded lack some of the information that the full patent documents that are in .pdf format contain. The full documents can contain

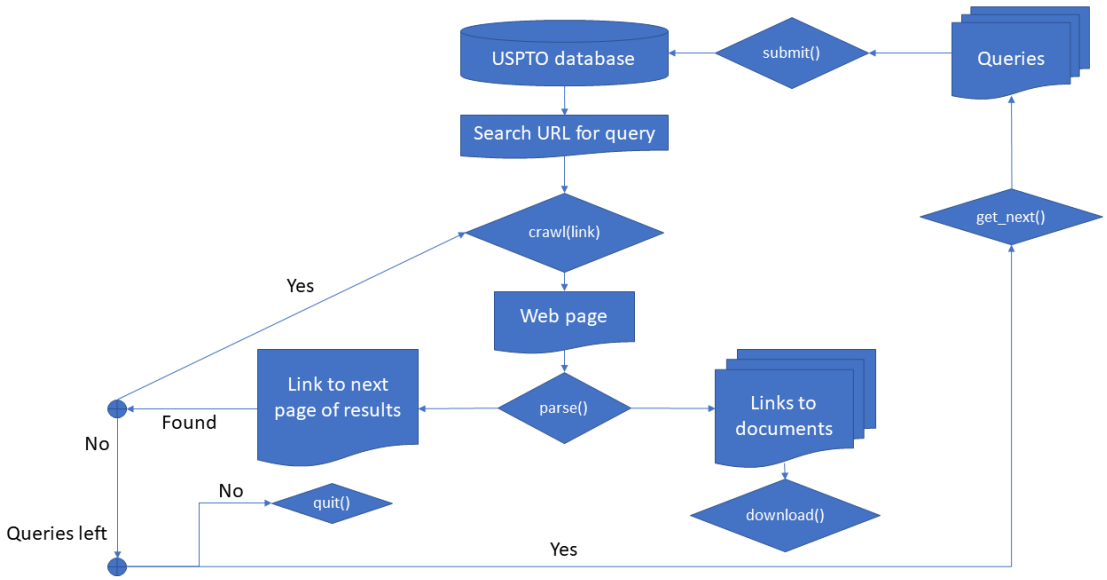

Figure 4.3. *Recursive patent document retrieval process.*

figures of designs, technical drawings, snippets of code or other such elements which give more visual depiction of the invention that is being claimed by the patent. For the sake of ease of processing the information contained within the patents and because of the format that the full documents are accessible in, only the .hmtl versions are downloaded, but the full contents of some of the documents are also examined in some cases. For these reasons, the term *patent document* is used in this thesis to refer to the .html contents rather than the full documents, if not specified otherwise.

Additional scripts were created to parse the downloaded documents to extract relevant information from them, such as filing and granting date, patent classes and the names of inventors, applicants and assignees. The scripts used are divided into two sets. *Miner* is responsible for submitting queries and downloading documents, whereas *parser* scripts process documents which have been stored locally, extract relevant information fields from them and combine the information from those fields into a single .csv file, which contains the desired data points from all the patents that were retrieved. These data points which are stored in a single file can then be used for different types for data analysis and visualization. The process flow of information retrieval by the miner part of the created scripts is visualized in Figure 4.3.

Creating a set of queries that make it possible to obtain a representative sample of patents related to a given technology is a challenging task. As pointed out by [61], patents are not classified in such precision that it would be possible to conduct a search for patents related to a single technology by querying by classifications. Additionally, the names of all the technologies that are related to the patent are not always mentioned in the title or the abstract of a patent. Some other challenges also occur while searching for patent

Query ID	Query	Matches
Q1	"temporal contrast vision sensor"	54
Q2	"event-based sensor"	48
Q3	"dynamic vision sensor"	184
Q4	"asynchronous time-based image sensor"	28
Q5	"event-based vision"	79

Table 4.1. *Queries used for patent document retrieval from USPTO database*

documents in the USPTO database. The search algorithm includes some limitations when searching for exact phrases. For example, if there is a document in the database that contains the phrase, but with additional stop words, which are some commonly used words in patent documents [62], such as *that*, *where* or *the*, those documents are counted as matches for the given query. Due to this behaviour by the search system, a patent document containing the phrase *"event where the camera..."* is included in the results when the search is conducted by using the query *"event camera"*. The database search interface does not provide an option to list only exact matches, where the stop words are dismissed. To combat the undesirable behaviour of the USPTO search algorithm regarding stop words, additional filtering is conducted to get rid of the documents that do not contain the exact phrase that was searched for, without stop words. The query "event camera" is not among the ones used in the document retrieval, since the term is generic by nature and can be used in other contexts not relating to the event camera technology. The queries that are used are limited to contain more technical terms, that are considered unique to the event camera technology. The queries are picked from terms that commonly occur in scientific literature about event cameras, and in those patent documents that were found by initially searching by broader terms.

Queries for a given term also give results where the text only appears in the literature references section, which means that a decision has to be made whether to include those results. In this case, those results are chosen to be included, the reasoning behind being that a patent which refers to a scientific paper with the query term in its title, is related to the specific query term in question. Those types of patents represent 30.9 percent of the retrieved documents. Several queries related to event camera technology were used, and they are displayed in Table 4.1. The number of matches on table is the amount of patent documents after filtering, meaning that the ones without exact matches are not included in it.

Some of the retrieved patent documents are matches for multiple queries in Table 4.1. Examining the overlap of query results can provide information on how different terms used in querying are related to each other. The percentage of overlap between queries is visible in Figure 4.4, where the values presented tell what percentage of the results

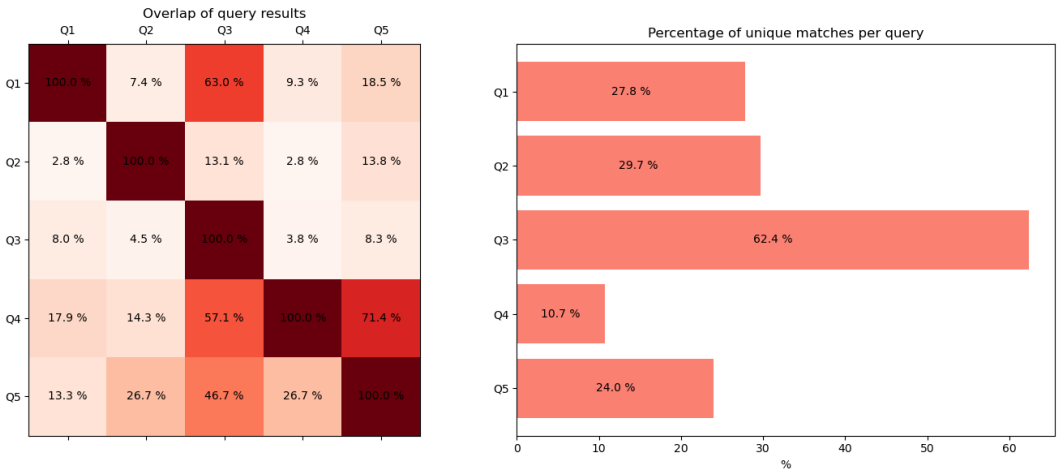

Figure 4.4. *Overlap of results between different queries presented in Table 4.1 and percentage of unique matches.*

of a given query in y-axis are also results for the query in x-axis. Percentage of unique matches for a given query, which means the percentage of patents matching a query that were not matches for any other query used, is also shown. Removing duplicate matches reduces the amount of patents retrieved significantly from initial expectations. If all the values from the column *Matches* in Table 4.1 are summed, the sum equals 396 matches. However, the real number of documents retrieved is 256 after dismissing duplicates from the results.

4.2.2 Patent data

Patent documents contains several elements which are required to understand before analyzing it. *Applicant* refers to the individual or organization which is responsible for filing the patent. The organizations capable of filing a patent application are narrowed down to legal entities [63]. *Assignee* is by definition the entity that has legal interests to the ownership rights of a patent. It is often the same organization that employs the inventor of the technology [64]. *Inventors* are the individuals who have contributed intellectual work towards making the patent contents possible. *References* in patents can point to either related patent documents or other literature, such as scientific papers. *Claims* define the scope of a patent from the perspective of what about a given patent is legally enforceable [63]. *Forward citations* of a patent are other patents which have been published after the first one, and have listed the older patent in their referred patents. Number of forward citations can be considered as a metric of commercial success of a patent, as noted by Henderson et al. [65], who ranked patents to winners and losers based on the number of forward citations that the patents had. For this reason, event camera patent data is searched for the top cited patents to find out which patents are the most successful in this regard. Top 10 patents by this metrics are presented in Table 4.2.

Patent number	Forward citations	Patent title	Applicants	Assignees
8780240	20	Method for the generation of an image in electronic form, picture element (pixel) for an image sensor for the generation of an image as well as image sensor	POSCH, CHRISTOPH; LITZENBERGER, MARTIN; MATOLIN, DANIEL; WOHLGENANNT, RAINER	AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH
7573956	19	Time encoding and decoding of a signal	Not specified	THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
9143680	15	Event-based image processing apparatus and method	SAMSUNG ELECTRONICS CO., LTD.	SAMSUNG ELECTRONICS CO., LTD.
9001220	10	Image sensor chip, method of obtaining image data based on a color sensor pixel and a motion sensor pixel in an image sensor chip, and system including the same	SAMSUNG ELECTRONICS CO., LTD.	SAMSUNG ELECTRONICS CO., LTD.
9471840	10	Apparatus and method for low-power object-detection in images using hardware scanning window	QUALCOMM INCORPORATED	QUALCOMM INCORPORATED
9554100	10	Low-power always-on face detection, tracking, recognition and/or analysis using events-based vision sensor	QUALCOMM INCORPORATED	QUALCOMM INCORPORATED
9804635	10	Electronic device and method for controlling displays	SAMSUNG ELECTRONICS CO., LTD.	SAMSUNG ELECTRONICS CO., LTD.
10032498	8	Memory cell unit and recurrent neural network including multiple memory cell units	SAMSUNG ELECTRONICS CO., LTD.; UNIVERSITAET ZUERICH	SAMSUNG ELECTRONICS CO., LTD.; UNIVERSITAET ZUERICH
10229341	7	Vector engine and methodologies using digital neuromorphic (NM) data	VOLKSWAGEN AG; AUDI AG; PORSCHE AG	VOLKSWAGEN AG; AUDI AG; PORSCHE AG
10133944	6	Digital neuromorphic (NM) sensor array, detector, engine and methodologies	VOLKSWAGEN AG; AUDI AG; PORSCHE AG	VOLKSWAGEN AG; AUDI AG; PORSCHE AG

Table 4.2. *Top 10 U.S. event camera patents by forward citations*

Table 4.2 shows that among the top patents by forward citations, most of the patents are related to the event sensors and core data processing, but patents granted to more advanced applications of the sensor technology have made it to the list as well. Those patents consider applications related to object detection and tracking, and notably two of those are from automotive manufacturers which implies that there is possibility of automotive integration of event camera technology.

To examine event camera technology from the TLC perspective, the patents that were retrieved are plotted by their granting date in Figure 4.5 in S-curve format. By looking at Figure 4.5 where the accumulated number of patents and the yearly increases to that number is displayed, it can be seen that the development of event camera technology which started in the late 80's, did not prove commercial potential until 2009 when the first patent was published, and the threshold of 50 patents in total was not broken until 2017. After that, patents granted have seen exponential growth, increasing by over five-fold from 2017 to 2020. Comparing the cumulative patents of event camera technology to the theoretical S-curve would suggest that the technology is currently either at late emergence or at early growth stage of the TLC. However, as Figure 4.5 shows, the number of granted patents in 2020 was slightly less than the number granted in 2019. This might suggest the beginning of an early saturation stage, but one sample is not enough to confidently assume that that is the case, and inspection of other metrics should be conducted.

Each patent is assigned to one or several classes and sub-classes that correspond to the International Patent Classification (IPC) scheme, which is maintained by World Intellectual Property Organization [66]. Examining which classes are dominant in the patent documents of a given technology gives some hint on what industries might contain early adopters in the case of upcoming technology or main researchers in already established one. Also, if patent data of a given technology shows that there are not many patents

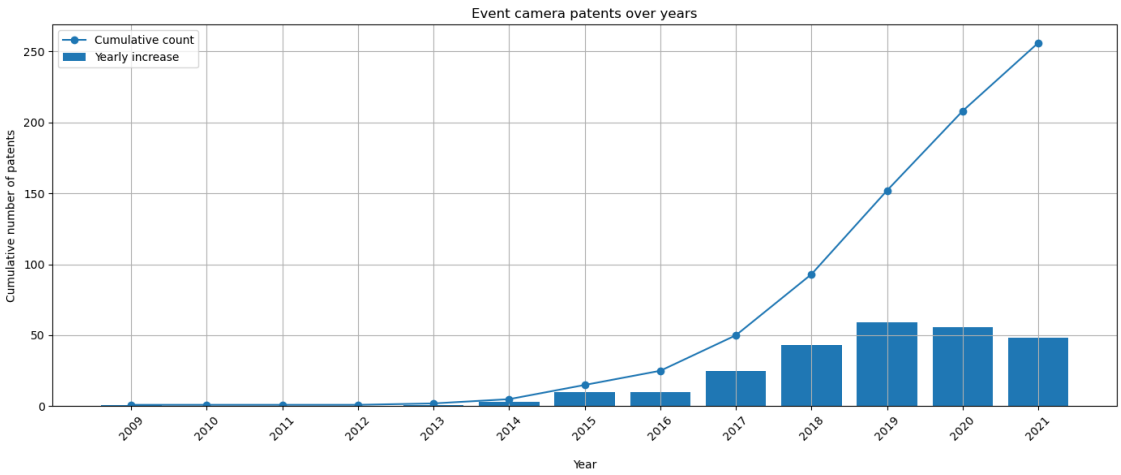

Figure 4.5. *Cumulative patents*

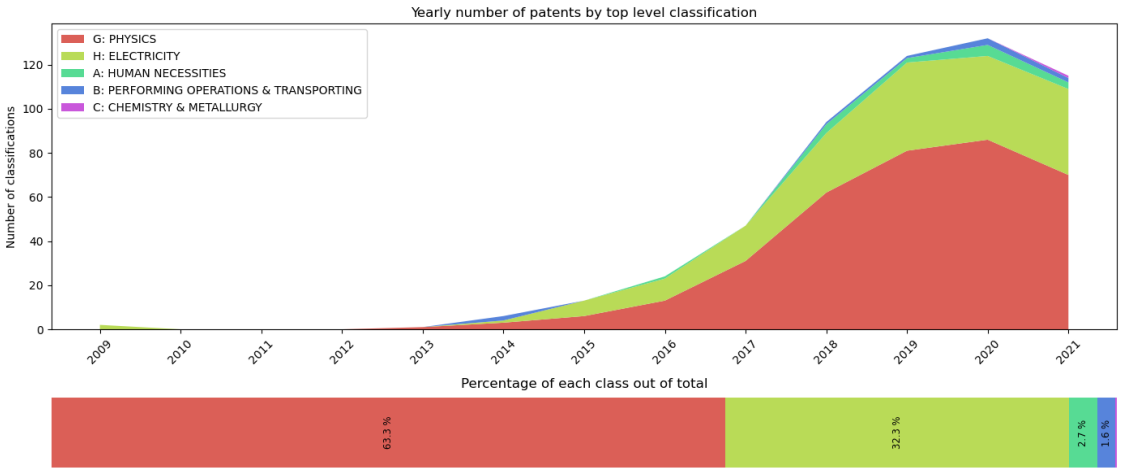

Figure 4.6. *Number and percentage of each top-level classification in patents by year*

outside the scope of classes that the technology is initially based upon, one can make an assumption that the technology does not yet have practical uses in wide range of fields. For example, in the case of event camera technology, from the classes of the patents, it can be seen in Figure 4.6 that majority of the patent classifications are made in core categories related to the technology itself, physics (G) and electricity (H). However, already 4.4 % of the classifications are in other categories, and with most of those types of patents being granted in the last three years. This hints that the event camera technology already has some practical usages that span into other industries. The potential for expansion towards other fields is examined with quantitative metrics in further sections of this chapter.

In addition to the top classes among the patents, sub-classes should be inspected also. Each sub-class that occurs in the patent data is displayed in Table 4.3 alongside the number of occurrences and a detailed description. Classifications show that among the patents in categories G and H, on which it is harder to determine the exact are of usage,

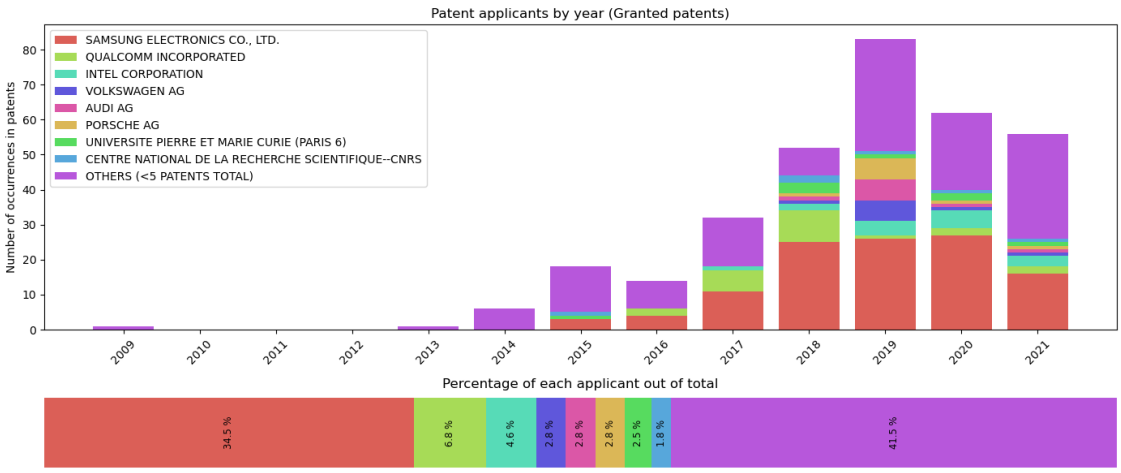

Figure 4.7. *Yearly applicants for granted event camera patents*

event camera technology has been granted sector-specific patents at least in the medical and automotive sectors.

One way to understand where a technology is possibly heading is to look which entities are being granted most patents. In this case, it will especially looked at the applicants to see if there are entities among them that are closely tied to the mobile phone industry. First, we can look at top applicants contributing to the technological development of event camera technology, meaning the applicants who have been granted most patents. Figure 4.7 shows the amount of patents from top applicants on a yearly basis.

Although Samsung is insurmountably the top applicant in patents related to event camera technology, it is worth noticing that as it was mentioned in chapter 2, the only product sold by them where they have used DVS sensors has been already pulled from the market. However, as was seen in chapter 2, the latest publication regarding sensor specifications by Samsung was published in 2020, which along with the patent data shows that the R&D efforts by the company continue. When considering the possibility of event camera integration into mobile phones, it is reasonable to assume that the knowledge that a company that operates in mobile phone sector have accumulated in other areas about any given technology is also utilized in mobile phones in future if that is perceived as useful. Two other enterprises, Sony and Apple, that have a stake in mobile imaging business are also found among the applicants, even though they are not among the top ones.

Qualcomm Incorporated, the applicant with the second largest number of granted patents on event camera technology, is the world's biggest supplier of smartphone application processors [67]. This brings some confidence to the argument that the event camera integration is being at least researched and considered in the mobile phone industry.

It is worth noticing that among the top 8 applicants are three companies that are focused

Class	Count	Description
A61B	8	DIAGNOSIS; SURGERY; IDENTIFICATION
A61F	2	FILTERS IMPLANTABLE INTO BLOOD VESSELS; PROSTHESES; DEVICES PROVIDING PATENCY TO, OR PREVENTING COLLAPSING OF, TUBULAR STRUCTURES OF THE BODY, E.G. STENTS; ORTHOPAEDIC, NURSING OR CONTRACEPTIVE DEVICES; FOMENTATION; TREATMENT OR PROTECTION OF EYES OR EARS; BANDAGES, DRESSINGS OR ABSORBENT PADS; FIRST-AID KITS
A61N	4	ELECTROTHERAPY; MAGNETOTHERAPY; RADIATION THERAPY; ULTRASOUND THERAPY
A63F	1	CARD, BOARD, OR ROULETTE GAMES; INDOOR GAMES USING SMALL MOVING PLAYING BODIES; GAMES NOT OTHERWISE PROVIDED FOR
B25J	2	MANIPULATORS; CHAMBERS PROVIDED WITH MANIPULATION DEVICES
B32B	2	LAYERED PRODUCTS, i.e. PRODUCTS BUILT-UP OF STRATA OF FLAT OR NON-FLAT, e.g. CELLULAR OR HONEYCOMB, FORM
B60R	1	VEHICLES, VEHICLE FITTINGS, OR VEHICLE PARTS, NOT OTHERWISE PROVIDED FOR
B60T	1	VEHICLE BRAKE CONTROL SYSTEMS OR PARTS THEREOF; BRAKE CONTROL SYSTEMS OR PARTS THEREOF, IN GENERAL
B60W	2	CONJOINT CONTROL OF VEHICLE SUB-UNITS OF DIFFERENT TYPE OR DIFFERENT FUNCTION; CONTROL SYSTEMS SPECIALLY ADAPTED FOR HYBRID VEHICLES; ROAD VEHICLE DRIVE CONTROL SYSTEMS FOR PURPOSES NOT RELATED TO THE CONTROL OF A PARTICULAR SUB-UNIT
B62D	1	MOTOR VEHICLES; TRAILERS
C12Q	1	MEASURING OR TESTING PROCESSES INVOLVING ENZYMES OR MICRO-ORGANISMS
G01B	9	MEASURING LENGTH, THICKNESS OR SIMILAR LINEAR DIMENSIONS; MEASURING ANGLES; MEASURING AREAS; MEASURING IRREGULARITIES OF SURFACES OR CONTOURS
G01C	3	MEASURING DISTANCES, LEVELS OR BEARINGS; SURVEYING; NAVIGATION; GYROSCOPIC INSTRUMENTS; PHOTOGRAMMETRY OR VIDEOGRAMMETRY
G01J	18	MEASUREMENT OF INTENSITY, VELOCITY, SPECTRAL CONTENT, POLARISATION, PHASE OR PULSE CHARACTERISTICS OF INFRARED, VISIBLE OR ULTRA-VIOLET LIGHT; COLORIMETRY; RADIATION PYROMETRY
G01M	1	TESTING STATIC OR DYNAMIC BALANCE OF MACHINES OR STRUCTURES; TESTING STRUCTURES OR APPARATUS NOT OTHERWISE PROVIDED FOR
G01N	2	INVESTIGATING OR ANALYSING MATERIALS BY DETERMINING THEIR CHEMICAL OR PHYSICAL PROPERTIES
G01P	2	MEASURING LINEAR OR ANGULAR SPEED, ACCELERATION, DECELERATION, OR SHOCK; INDICATING PRESENCE, ABSENCE, OR DIRECTION, OF MOVEMENT
G01S	11	RADIO DIRECTION-FINDING; RADIO NAVIGATION; DETERMINING DISTANCE OR VELOCITY BY USE OF RADIO WAVES; LOCATING OR PRESENCE-DETECTING BY USE OF THE REFLECTION OR RERADIATION OF RADIO WAVES; ANALOGOUS ARRANGEMENTS USING OTHER WAVES
G02B	12	OPTICAL ELEMENTS, SYSTEMS, OR APPARATUS
G02C	1	SPECTACLES; SUNGLASSES OR GOGGLES INSOFAR AS THEY HAVE THE SAME FEATURES AS SPECTACLES; CONTACT LENSES
G03B	5	APPARATUS OR ARRANGEMENTS FOR TAKING PHOTOGRAPHS OR FOR PROJECTING OR VIEWING THEM; APPARATUS OR ARRANGEMENTS EMPLOYING ANALOGOUS TECHNIQUES USING WAVES OTHER THAN OPTICAL WAVES; ACCESSORIES THEREFOR
G05D	4	SYSTEMS FOR CONTROLLING OR REGULATING NON-ELECTRIC VARIABLES
G05F	1	SYSTEMS FOR REGULATING ELECTRIC OR MAGNETIC VARIABLES
G06F	85	ELECTRIC DIGITAL DATA PROCESSING
G06K	77	RECOGNITION OF DATA; PRESENTATION OF DATA; RECORD CARRIERS; HANDLING RECORD CARRIERS
G06N	12	COMPUTER SYSTEMS BASED ON SPECIFIC COMPUTATIONAL MODELS
G06Q	1	DATA PROCESSING SYSTEMS OR METHODS, SPECIALLY ADAPTED FOR ADMINISTRATIVE, COMMERCIAL, FINANCIAL, MANAGERIAL, SUPERVISORY OR FORECASTING PURPOSES; SYSTEMS OR METHODS SPECIALLY ADAPTED FOR ADMINISTRATIVE, COMMERCIAL, FINANCIAL, MANAGERIAL, SUPERVISORY OR FORECASTING PURPOSES, NOT OTHERWISE PROVIDED FOR
G06T	80	IMAGE DATA PROCESSING OR GENERATION, IN GENERAL
G08B	7	SIGNALLING OR CALLING SYSTEMS; ORDER TELEGRAPHS; ALARM SYSTEMS
G08G	1	TRAFFIC CONTROL SYSTEMS
G09B	1	EDUCATIONAL OR DEMONSTRATION APPLIANCES; APPLIANCES FOR TEACHING, OR COMMUNICATING WITH, THE BLIND, DEAF OR MUTE; MODELS; PLANETARIA; GLOBES; MAPS; DIAGRAMS
G09G	14	ARRANGEMENTS OR CIRCUITS FOR CONTROL OF INDICATING DEVICES USING STATIC MEANS TO PRESENT VARIABLE INFORMATION
G10K	1	SOUND-PRODUCING DEVICES
G10L	2	SPEECH ANALYSIS OR SYNTHESIS; SPEECH RECOGNITION
G11C	3	STATIC STORES
H01L	19	SEMICONDUCTOR DEVICES; ELECTRIC SOLID STATE DEVICES NOT OTHERWISE PROVIDED FOR
H03D	1	DEMODULATION OR TRANSFERENCE OF MODULATION FROM ONE CARRIER TO ANOTHER
H03F	5	AMPLIFIERS
H03K	4	PULSE TECHNIQUE
H03M	1	CODING, DECODING OR CODE CONVERSION, IN GENERAL
H04B	2	TRANSMISSION
H04J	2	MULTIPLEX COMMUNICATION
H04L	5	TRANSMISSION OF DIGITAL INFORMATION, e.g. TELEGRAPHIC COMMUNICATION
H04M	2	TELEPHONIC COMMUNICATION
H04N	129	PICTORIAL COMMUNICATION, e.g. TELEVISION
H04Q	1	SELECTING
H04R	1	LOUDSPEAKERS, MICROPHONES, GRAMOPHONE PICK-UPS OR LIKE ACOUSTIC ELECTROMECHANICAL TRANSDUCERS; DEAF-AID SETS; PUBLIC ADDRESS SYSTEMS
H04S	1	STEREOPHONIC SYSTEMS
H04W	5	WIRELESS COMMUNICATION NETWORKS
H05B	2	ELECTRIC HEATING; ELECTRIC LIGHTING NOT OTHERWISE PROVIDED FOR

Table 4.3. *IPC codes present in the event camera patent data*

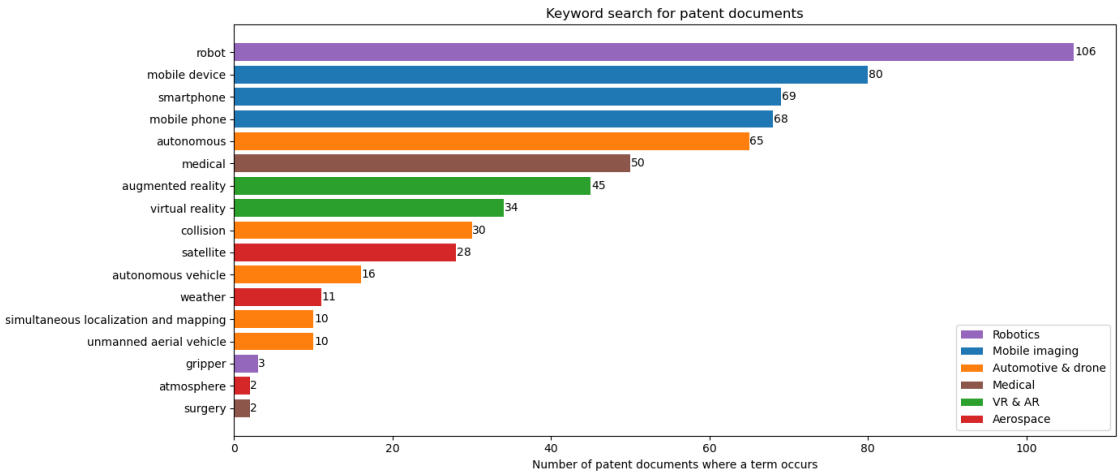

Figure 4.8. *Keyword term occurrences in event camera patent data*

on automotive industry, Volkswagen AG, Audi AG and Porsche AG, although the latter two are subsidiaries of the first one. Closer look into the patent documents also reveal that for all of the patents where one of these companies is listed as applicant, also one or two of the others are listed as applicant as well. For other automotive manufacturers, at least Honda's R&D vision has been granted one patent related to event camera.

Looking at the titles and contents of the patents suggests that there are several areas where the patent applications focus. Virtual- and augmented realities (VR and AR) are mentioned in several patents, along with areas such as mobile phones, automotive, robotics and aerospace. One way to look at the representation of different areas in the patent documents is by conducting a keyword search. Different keyword terms related to several industries are gathered and searched for in the patent documents that were retrieved. The number of documents containing each term is displayed in Figure 4.8.

Although mobile imaging category yields more keyword term matches from the patent data than other categories that were tested, closer examination of the patents containing those keywords shows that in many cases, mobile phones are mentioned just as one example of an applications where the technology described in the patents could be utilized, as the patents are not especially granted for mobile phone implementation of the innovations.

In addition to analyzing the number of occurrences of keywords, their importance in the texts can be analyzed as well. This can be done by measuring the *term frequency-inverse document frequency* (tf-idf) score of the keywords. The calculation for this metric [68] is presented in Equations 4.1, 4.2 and 4.3.

$$tf(w,d) = \frac{w(d)}{|d|} \tag{4.1}$$

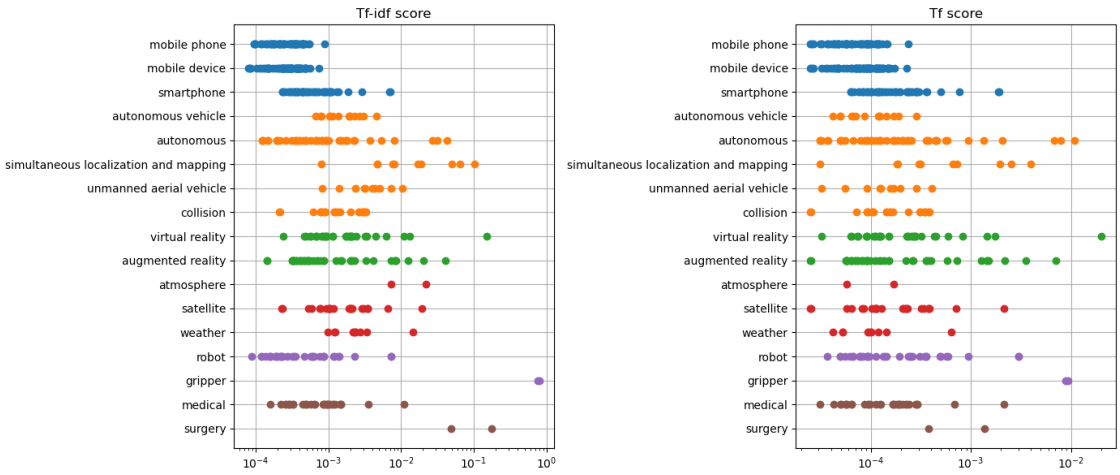

Figure 4.9. *Term frequency scores by themselves and with inverse document frequency*

$$idf(w, D) = 1 + log(\frac{N}{\sum_{i=1}^{N} W(w, d_i)}) \tag{4.2}$$

$$tf\text{-}idf(w, d, D) = tf(w, d) \times idf(w, D) \tag{4.3}$$

In Equations 4.1, 4.2 and 4.3, w is used to note the term for which the metric is being calculated for, d is a single document containing $|d|$ terms in total, and D is a set of N documents. Function $w(d)$ returns the amount of occurrences of the term w in a the document d, and $W(w, d)$ returns either 1 or 0 depending on if the document d contains the term w or not, respectively. The metric is therefore calculated for each document separately, but with respect to the complete set of documents. Tf and tf-idf scores for the same keyword terms that were used in Figure 4.8 are presented in Figure 4.9.

Using tf-idf scores can be problematic in cases where there exists an abbreviation for the term which is used in the text after only mentioning the full term once. The keyword list used here contains several terms like that, which include simultaneous localization and mapping (SLAM), unmanned aerial vehicle (UAV) and virtual and augmented realities (VR & AR). For those terms, the tf scores of abbreviations and the original terms are added together 4.9.

4.3 Forecasting and stage determination

Determining whether a technology is going to succeed or fail is a difficult task, but one that investors must consider every time they allocate their resources into any new or existing technology. Several studies have been done with different types of metrics and methods to gain foresight on how technologies may perform in the future. Here, some of those metrics will be utilized in order to determine current status and future prospects

of event camera technology. As it was found out previously by examining the patent data, the scope of event camera technology is not confined to contain only the hardware (event sensors) but also includes the applications that utilize that hardware for different purposes.

By looking at the S-curve of cumulative patents in Figure 4.5, initial assumption can be made that event camera technology is currently at the growth stage, but the data point representing the latest full year of observations (2020) does show signs of decreasing growth, which should not be dismissed. Altuntas et al. [52] conducted a study where they examined three different metrics driven from patent data to determine the future prospects of three different technologies, TFT-LCD displays, flash memory systems and personal digital assistants (PDA's). Motivation behind the study was to be able to analyze the investment potential of different technologies that are in the growth stage of the TLC, as is the event camera technology, according to the initial assumption. The metrics that they used were diffusion speed, expansion potential and patent power. The data querying criteria in their study was more strict than used in this thesis, and only patents with the specific term in the patent document's title were approved. For that reason, when calculating those metrics for event camera technology, two values will be calculated, one with the requirement that the query term is on the title, and one with the all the data gathered with the requirements mentioned previously. However, in the original study, no arguments were presented for why only the title matches were required. Perhaps the reason was that it resulted in smaller sample size, since the technologies that were analyzed would yield thousands of patents as result if all the patents containing the query term would be taken into account. As stated by [69], headlines can often be formulated in a way that the technology itself is not mentioned in it.

As for the metrics, *diffusion speed* originally introduced by Huang & Wang [70] looks at the value of the patents from the perspective of forward citations, giving patents that get cited more value in same manner as it was done earlier in this chapter in Table 4.2. Diffusion speed is defined as the total number of forward citations in the patent data divided by the total number of patents. Patents that have been granted within last four years from the time that querying happens are discarded when calculating diffusion speed, due to the fact that patents may take several years to undergo the examination period before they are granted, and therefore it is not reasonable to assume that patents citing other recent patents have been granted yet. In the case of event camera, patents from the last four years equal most of the patents that were retrieved, which results in a small sample size as can be seen from Figure 4.5. *Expansion potential* is defined as the total number of different IPC classifications that are present in the patent data, and *patent power* is the expansion potential divided by the total number of patents.

The study by [52] was published in 2014 and the patent data they used was gathered from the years prior to 2012, which means that the metrics that they calculated are now

Metric	TFT-LCD (-2012)	Flash memory system (-2012)	Personal digital assistant (-2012)	Event camera technology (title match required)	Event Camera technology (all)
Diffusion speed	10.61	31.11	31.46	6.50	3.42
Expansion potential	21	21	52	206	558
Patent power	1.73	1.30	1.38	2.30	2.12

Table 4.4. *Metrics used and values calculated by Altuntas et al. [52] and corresponding values calculated for event camera technology*

outdated. However, this makes it possible to compare the metrics they calculated with the actual outcomes of these technologies in the years that followed. The values for the three metrics calculated in the original study and calculated for event camera technology by two different criteria are displayed in Table 4.4.

Notable differences between the technologies compared is that the diffusion speed is considerably lower with event camera technologies than with any of the technologies in the original study, and that the patent power and expansion speed are both higher than with any of the other technologies in comparison. Higher expansion potential is in line with the fact that as listed in Chapter 2 and based on examining characteristics of the retrieved patent data, there exists many possibilities for event camera use cases in different industries.

The three technologies considered by [52], the metrics calculated for them and the resulted outcomes provide interesting outlook from the TLC perspective. TFT-LCD displays can still be considered prevalent, as the annual units shipped in 2019 was 770.5 million, but that is only 2.17 % increase from 2012 [71]. As the demand for other types of displays, such as OLED an AMOLED, have soared during recent years [71], while the increase in TFT-LCD display sales have been only modest, it is reasonable to consider that TFT-LCD technology has possibly reached the saturation stage of it's TLC. The argument gains more confidence when by looking at the S-curve of TFT-LCD technology in Figure 4.10, where the outcomes for the three technologies is displayed. Figure 4.10 shows that the growth of accumulated number of granted patents has decreased to near zero, a clear sign of saturation stage in the S-curve model of TLC.

The case of personal digital assistants depicts the scenario on which a technology goes through dramatic shift and moves on to a completely new S-curve. Even though smartphones have made PDA devices almost obsolete in the last decade, it can be argued that the smartphone technology was built on top of the base that PDA devices had established, or even that smartphones are actually a part of the PDA technology. If smartphones were to be considered as successors of PDA's or would be considered as part of the PDA technology, that would be in line with the fact that PDA technology was assigned the highest expansion potential of the three technologies compared. The relationship between PDA's and smartphones could be studied further by collecting patents related to smartphones and examining how often they cite PDA patents.

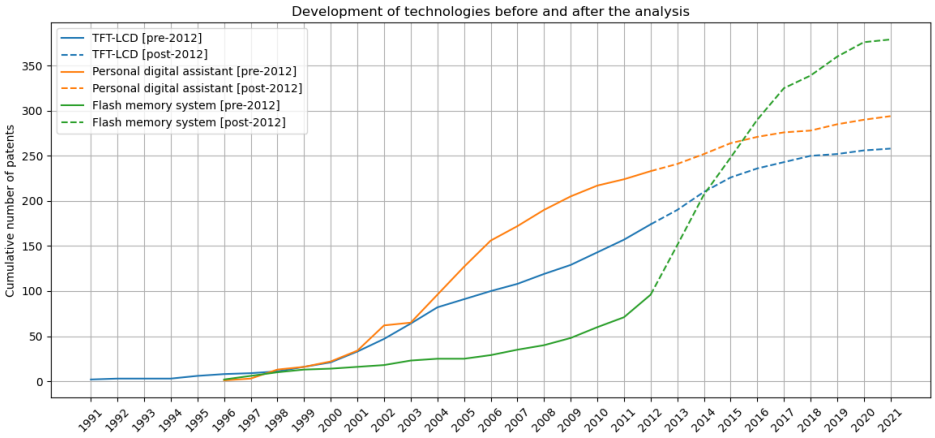

Figure 4.10. *S-curves for TFT-LCD, PDA and flash memory system technologies before and after the analysis by [52]*

As previously stated, the question that these metrics are supposed to help to answer is whether a technology is worth investing. If we look at the market data of the products in question, flash memory systems and PDA's saw large-scale growth, if PDA's are classified as smartphone predecessors [72] [2], whereas TFT-LCD has seen competing technologies gaining more interest and only modest growth over the years [71]. Obviously there exists other factors that affect the profitability of an investment in a technology, but growth in market size can be considered as a generally useful metric. When comparing only these three samples and the values calculated for them in Table 4.4 to the market size growth, it would seem that larger diffusion speed and lower patent power result in growth, whereas expansion potential cannot be used to predict future behaviour. That would suggest that event camera technology is not a viable investment target. However, the sample size of three is too small to suggest that any conclusions about the predictability of a technology's success could be made based on these metrics.

4.3.1 S-Curve models

Liu and Wang introduced a straightforward method to utilize three existing metrics for TLC forecasting, based on three different but similar equations used in modeling the S-curve [73]. The models in the study include the Loglet Lab model, the Pearl model and the Gompertz model. The Loglet Lab model is more commonly referred to as *logistic model*, for example by [50] and [74], and will for that reason be referred to as logistic model in this thesis.

The Logistic and Pearl models are better at describing technologies that experience a rapid growth phase [73], which makes them suitable candidates in the case of event cameras, for which there has been increasing growth until the year 2019, like seen in Figure 4.5. Contrary, the Gompertz model is better for modeling technologies that grow

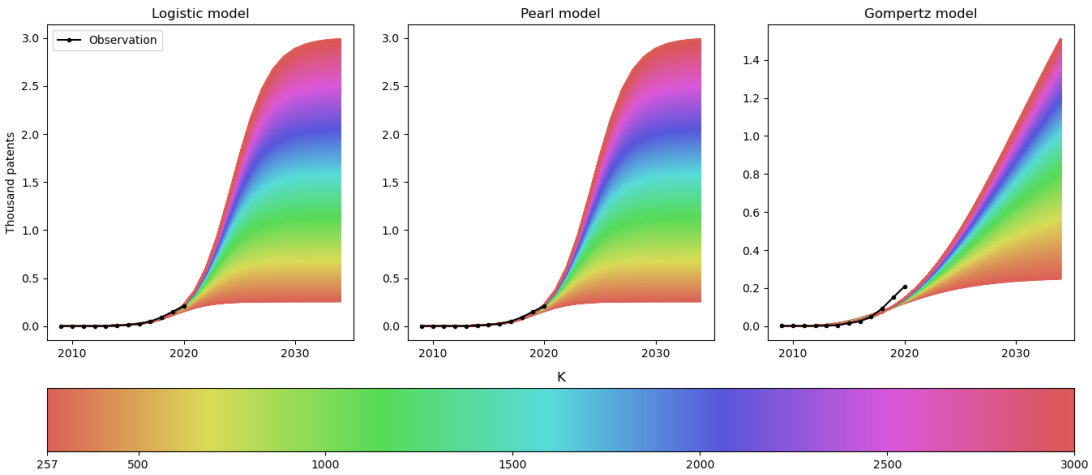

Figure 4.11. *Observations and forecasts for event camera technology by Logistic, Pearl and Gompertz S-curve models, using a range of different values for* K

with slower pace. The formulas for the S-curves that Logistic, Pearl and Gompertz models are described as follows:

$$Y_t = \frac{K}{1 + e^{-a(t-b)}}, \quad \text{Logistic model} \tag{4.4}$$

$$Y_t = \frac{K}{1 + ae^{-bt}}, \quad \text{Pearl model} \tag{4.5}$$

$$Y_t = Ke^{-ae^{-bt}}, \quad \text{Gompertz model} \tag{4.6}$$

In Equations 4.4, 4.5 and 4.6, Y_t is the value of the S-curve, K is the accumulated number of patents after saturation, t is the time point for which the value of Y_t is calculated, and a and b are the model parameters. The model equations are used to fit the existing real data and use that to obtain a and b, and then using these parameters again with Equations 4.4, 4.5 and 4.6 for forecasting. Forecasts by these three models and with a range of K spanning from 257, which is one more than the number of patents retrieved, and up to 3 000 are represented in Figure 4.11. The year 2021 is dismissed when calculating the parameters, since at the time of the writing the year has not reached it's end, and additional patents will almost certainly be granted still. As previously mentioned, Gompertz model is suitable for technologies with more moderate growth pace, and in Figure 4.11 that is evident, as the Gompertz model forecasts separate from the real observations early on. For this reason, the Gompertz model is not addressed anymore in the follow-up analysis of the models. Logistic and Pearl models also provide identical values for the S-curve, and for that reason only the logistic model will be used from now on.

In the study by Liu and Wang, determination of the best value for K was done by com-

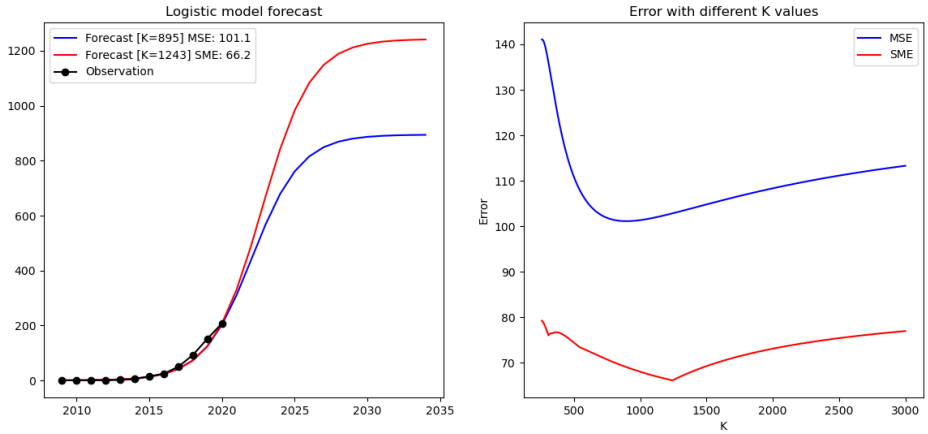

Figure 4.12. *Optimal forecast curves chosen by calculating minimum MSE and SME by trying different values of K. On the right, corresponding error values are displayed.*

puting forecasts by using a range of different values for K, like in Figure 4.11, and then calculating MSE and the Sum of Modulus Error (SME), also known as the sum of absolute errors (SAE), between the forecast and the observation curves, and then choosing as best value the K that resulted in minimum value for both of these error metrics. SME calculation for arrays X and Y which are of length N is visible in Equation 4.7, and the forecast curves that correspond to the value of K causing the least error by both of these metrics is visualized in Figure 4.12.

$$SME(X,Y) = SAE(X,Y) = \sum_{i=0}^{N} |X_i - Y_i| \tag{4.7}$$

The model forecasts that the number of event camera patents will grow rapidly until the late 2020's, and that the saturation value is between 895 (minimum MSE) and 1243 (minimum SME), which would indicate that the total number of patents would increase from 3.50 to 4.86 times the current amount. If we look at the theoretical S-curve presented in Figure 4.1, we can see that the transition from emergence to growth stage happens a short period before the rate of growth in cumulative patents reaches it's maximum value. Using this notion and looking at the forecast S-curve and the position of latest observations related to it, it can be seen that the point on the forecast curve at the position of latest observation is already at the highest rate of growth, but the sequence of latest observed points show growth at slightly slower rate. This suggests that the event camera technology is either at the verge of transition from emergence to growth, or has recently passed the transition point.

The patent data for other technologies that was gathered previously can be utilized to validate the functioning of these S-curve models. These technologies are especially useful, since they have all arguably reached the saturation stage based on the S-curves,

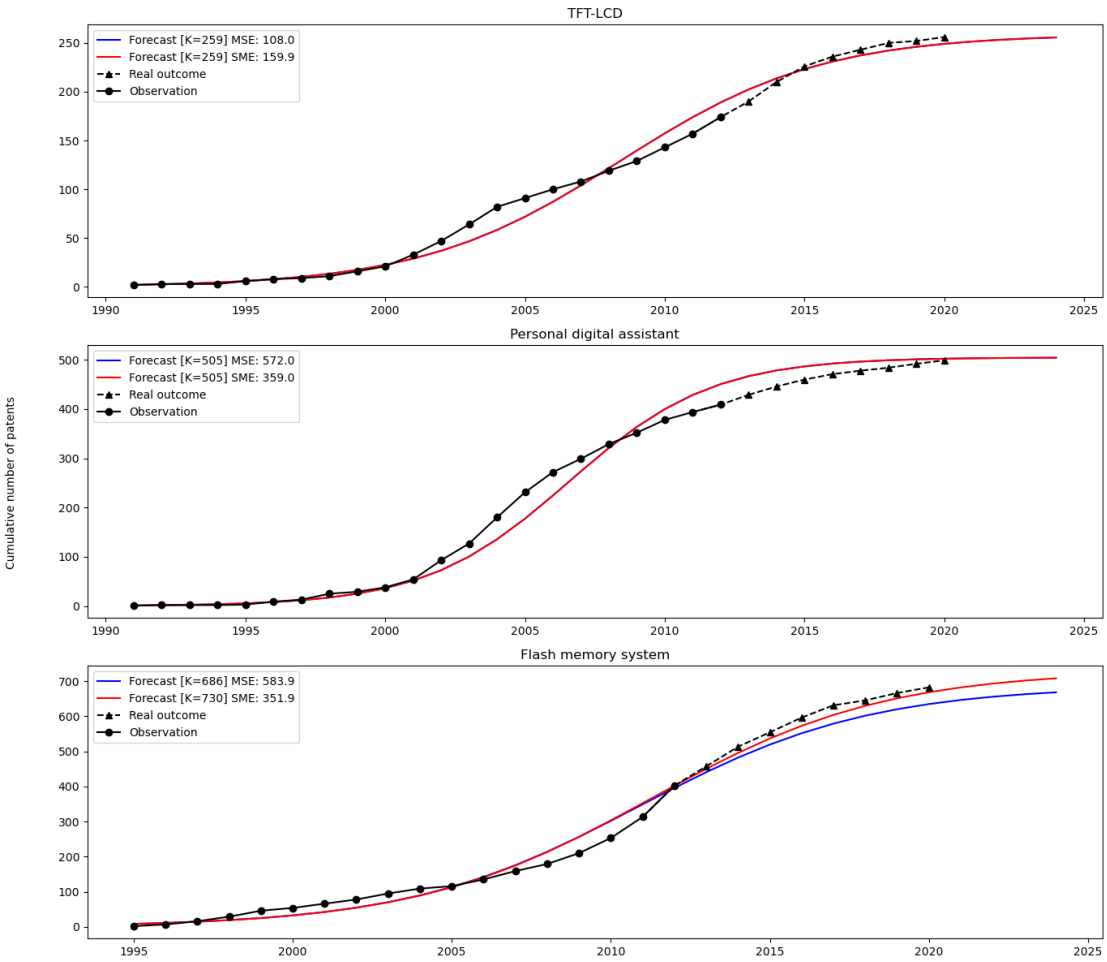

Figure 4.13. *Forecasts using logistic model and patent data of TFT-LCD, PDA and flash memory system technologies prior to 2012, compared with the actual outcomes that occurred*

and therefore they are considered suitable candidates for examining how accurate the error-based fitting and determining of K is. These technologies are not however the most ideal for this usage, since they are not part of the same product family which would be desirable for comparison purposes as pointed out by [50]. However, patent data for such technologies which would show clear signs of saturation on the S-curve in the imaging sensor space was not able to be found. Data from these three technologies prior to 2012 is fitted and compared to the actual outcomes in the accumulated patent numbers that followed after that year. The forecasts compared to observations are presented in Figure 4.13.

The logistic model is able to predict the development of these three technologies in high accuracy. Perhaps the most notable of these is the forecast concerning flash memory systems. In that case, latest years of observations used for forecast construction showed accelerating growth, but nevertheless the model was able to accurately predict that the growth will slow down. The observations used for constructing the forecasts contain more

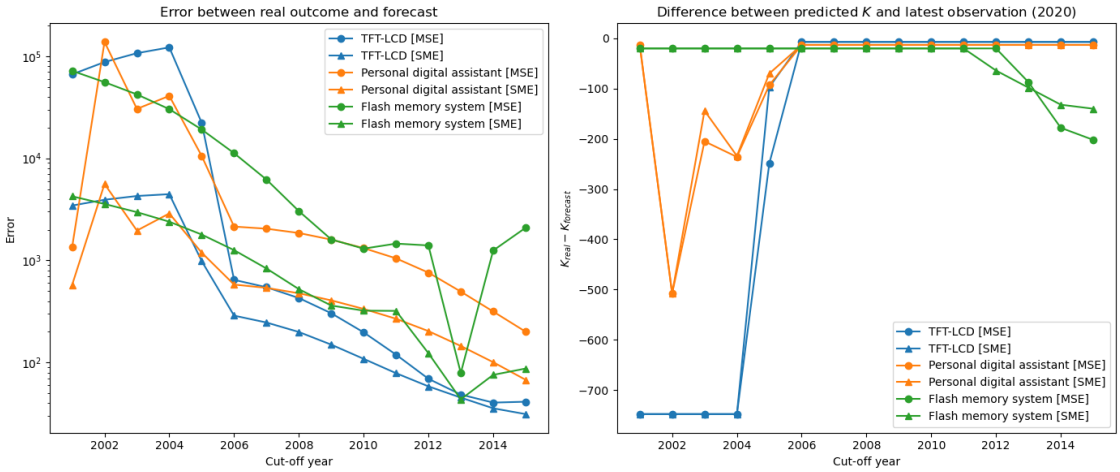

Figure 4.14. *Forecast accuracy with different amount of data points available. Cut-off year in x-axis means the latest year from which data is used as input for the forecast model.*

data points than the observations used for event camera forecasts, exactly 22 in the case of TFT-LCD and PDA, and 18 in the case of flash memory systems, whereas event camera technology patent data has only 12 data points. For this reason, it should be tested how far back the cut-off point where the data is split into observation and outcome can be moved, until the forecasts become unreliable. That test is visualized in Figure 4.14, where K_{real} equals the latest observation which is the number of patents that was granted by the end of the year 2020.

It can be seen from Figure 4.14 that the cut-off point can be rolled back all the way to the year 2006 until the difference between latest observed cumulative number of patents and predicted K starts to differ significantly, reducing the number of data points to 16 with TFT-LCD and PDA, and 12 with flash memory systems. Same cut-off year applies for the calculated MSE and SME between the real outcome and the forecasts. What is significant about the result of this test is that if we look at the S-curves of the three technologies and the visualized errors, we can see that if the data used for making the forecast has reached the point of rapid growth, the forecasts are reliable in these cases. It should be noted that the latest observed value of K, which is considered in the forecast accuracy determination to be the real outcome value, is actually slightly lower than the real saturation value which is yet to be determined, since although the growth in accumulated patents have been decreased to a low level, it is nevertheless still growing slightly every year.

One trait of the forecast models should be noted after looking at Figure 4.14, and that is the volatility of the forecast when having few data points. If we look at the difference between latest observation and predicted K in the case of TFT-LCD patents, it can be seen that the difference drops from near zero down to below -700 within just two-year

difference in the cut-off year, from 2006 to 2004. As the real observed K is around 250, the error of predicted K is over 150 % of the actual outcome value. Similar behaviour can also be observed within the PDA forecasts. This somewhat increases the doubts when considering the event camera forecast.

4.3.2 Entropy model

Entropy is a term that is used in thermodynamics to describe the amount of disorder in a system. Lin et al. refined the term to be used in the area of TLC examination [53]. In the TLC context, the larger the diversity is among patent applicants in the technology space, the larger the entropy value becomes. On the contrary, if most of the patents are applied by only few applicants, the entropy value is smaller. Lin et al. also examined how the value behaves in different TLC stages and in transitions between them, and found out some trends that occur in those scenarios. The basis for the calculation of the entropy indicator for TLC analysis is presented in Equation 4.8.

$$H = -\sum_{i=1}^{M} \frac{N_i}{\sum_{j=1}^{M} N_j} log \frac{N_i}{\sum_{j=1}^{M} N_j} \tag{4.8}$$

In equation 4.8, H is the entropy value, M is the total number of patent applicants, N_i is the total number of patents granted and N_j is the number of patents by a single applicant. The indicator is calculated on a yearly basis.

In their analysis, Lin et al. noted that growth, the second of four TLC stages previously introduced is better to be divided into two distinct stages in the case of some technologies. They call these two stages preliminary and real growth, where preliminary growth precedes real growth and can occasionally experience more fluctuations than it is typical for the growth stage. From the entropy indicator data, one can estimate the current TLC stage that the technology is currently experiencing, and also transition points from one stage to other. Entropy calculated by using Equation 4.8 for the event camera patent data is presented in Figure 4.15.

In addition to the entropy calculated using applicant information, Figure 4.15 also shows the value calculated using assignee data, since those data points represent similar type of information, answering the question of who is behind the patent. According to Lin et al. the transition from emergence to preliminary growth stage happens after the value of H has risen from zero, when $\frac{dH}{dt} = 0$ and when H decreases after the transition. Based on the indicator values over time, it appears that the best candidate for transition point from emergence to growth is at year 2015, where the entropy reaches the condition $\frac{dH}{dt} = 0$ and decreases for one year until it starts to grow steadily when looking at the values calculated by assignee data. When looking at the applicant data, there is not considerable

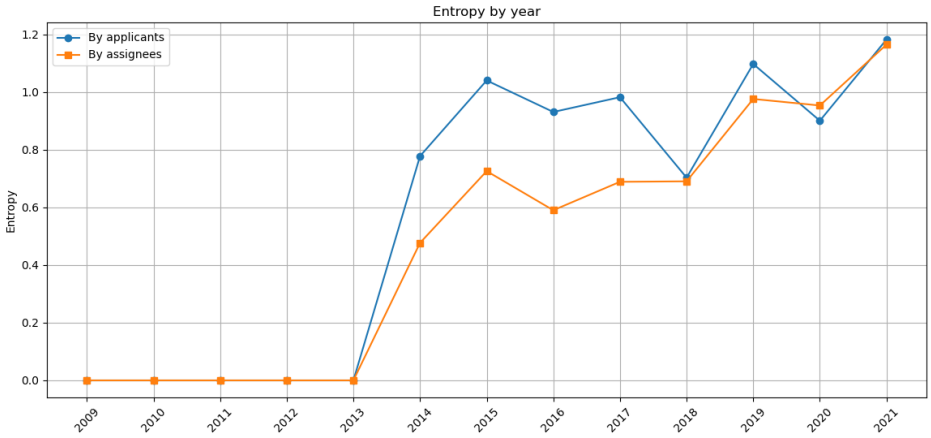

Figure 4.15. *Entropy indicator derived from event camera patent data*

growth after 2015, although the trend is going upwards with some fluctuatios. According to the entropy model, event camera technology is therefore possibly transitioned from emergence to growth phase, but other indicators must be observed to confirm or dismiss whether it has really happened.

4.3.3 Other indicators

Several other indicators can be derived from the patent data. Haupt et al. conducted a study on how those indicators behave in TLC stage transitions [61]. The indicators include the *average number of backwards citations* made to other patents and scientific papers, and their hypothesis was that the amount of citations of both types increase in transition from emergence to growth phase. *Average immediacy of patent citations* means the time that has passed between the cited patents and the patents that have cited them. Average number of forward citations, a term defined in earlier section of this chapter, is also included. *Average duration of examination process* is the time that has passed between filing and granting the patents. *Average number of claims* and *average number of priorities* are also considered. The study by Haupt et al. was conducted on pacemaker technology, which contains some different traits when comparing it to event camera technology. For example, pacemakers are heavily contained in the medical field, meaning that the number of applications outside of the scope of medical is limited. Summary of these indicators, the hypotheses by Haupt et al. and their observed results are visible in Table 4.5. Values for these indicators derived from event camera patent data are displayed in Figure 4.16.

The values of regarding forward citations might be misleading, because as Figure 4.16 shows, the duration of examination has ranged from around one up to six years in recent years. This means that some of the patents that will refer to the patents published recently

Indicator	Hypothesis	Observation by Haupt et al. [61]	Observation on event camera patent data
Backwards citations	Significant increase from emergence to growth and from growth to maturity in both patent and literature citations	Patent citations increased at both transitions, literature citations only from emergence to growth	Patent citations increased and literature citations decreased to a steady level
Immediacy of citations	Significantly higher in emergence and maturity stages than in growth stage	Increase was only observed when comparing growth to maturity stage	Slight drop after 2014
Forward citations	Significant increase from emergence to growth	Observations confirmed the hypothesis	Not conclusive
Claims	Increasing in every TLC stage	Increase was only observed when comparing growth to maturity stage	Steady throughout the years
Priorities	Increasing in every TLC stage	Increase was only observed when comparing growth to maturity stage	Increased to a steady level in 2014
Duration of examination	Significantly higher in emergence and maturity stages than in growth stage	Observations confirmed the hypothesis	Trend going downwards from 2013

Table 4.5. *TLC indicators and corresponding hypotheses by [61] and observations using event camera patent data*

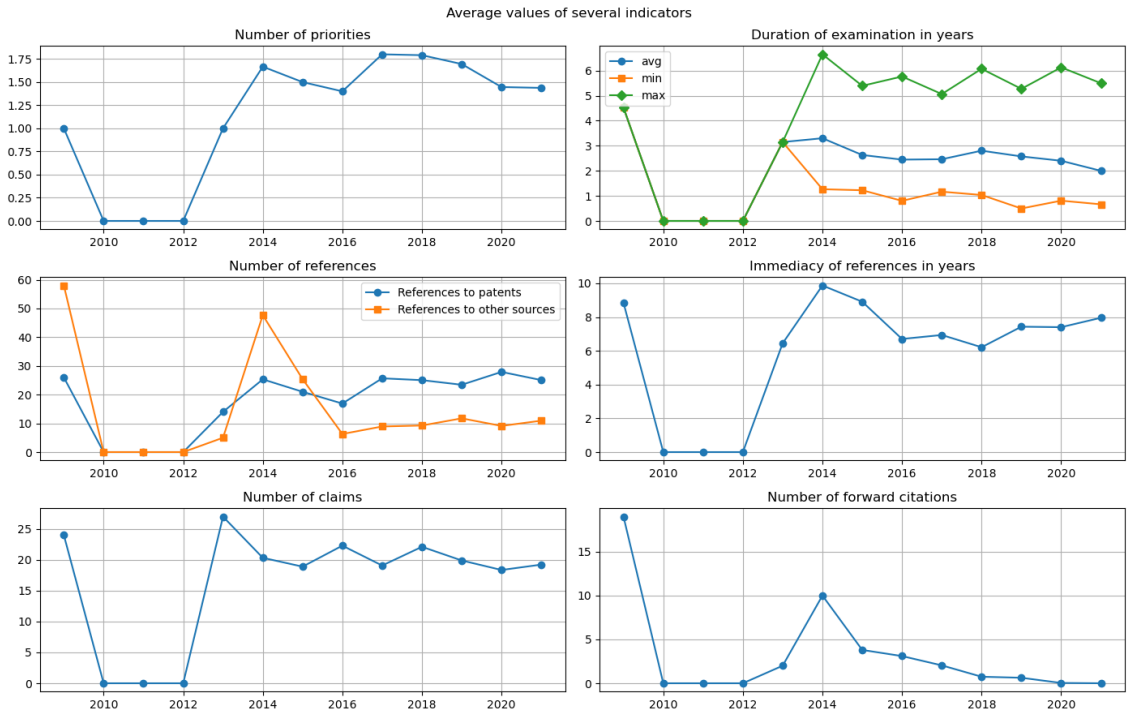

Figure 4.16. *TLC indicators derived from event camera patent data*

are yet to be published. Haupt et al. emphasize that in order to reliably address that a technology has transitioned between TLC stages, one must be able to observe a change in the indicators that is both significant and has a long-term effect [61]. Although it seems that most of the indicators increase significantly from 2012 to 2014, this effect should be observed with some criticism, since there were no granted patents related to event camera technology between 2010 and 2012, and the indicators have remained relatively close to their values in 2009, which should also not be given too much emphasis, since there was only one patent granted that year. For these reasons, it is not possible to make convincing arguments about the stage of event camera TLC based on these metrics.

Patent data offers useful insights into the development of technologies, but it can be misleading to rely solely on it when conducting TLC analysis. As Järvenpää et al. [69] point

out, other non-patent based indicators can also provide insights, and the way that these indicators behave is related to the stage of the life cycle. In their study the hypothesis was that in the early stages of the life cycle, scientific literature appears before other indicators. Scientific literature development can be measured by examining the Scientific Citation Index Expanded (SCIE), which measures the number of citations to papers related to a technology among a curated collection of 9 200 established journals. Scientific literature was thought to be followed by engineering papers, patent documents and news articles in respective order. The explanation behind the order of the indicators was that scientific literature depicts status of fundamental research that is required before applied research can occur, and which can in turn be measured by indicators derived from engineering paper publishing data. Patent documents come to picture only on when further development happens and commercial applicability is proven possible, and the actual inflection of a technology can be seen from examining news articles related to the technology. However, the hypothesis which was tested by three different technologies was concluded to be not generally applicable principle, as in the case of the tested technologies examined the order was different.

The non-patent indicators were calculated for event camera technology using the same query terms that were used in the patent search, and which were presented in Table 4.1. In the case of literature about event camera technology, it is very difficult to make a distinction on whether a publication is part of the fundamental scientific research or applied engineering, and therefore the two metrics are combined, and citation index value thus contains the citations to all the papers that match the queries used. News articles matching the queries are obtained by using the Nexis Uni search engine [75], which offers matches from over 17 000 different publications. Citation indices and the number of occurrences of the query terms in news articles is presented in Figure 4.17. Query Q4 (asynchronous time-based image sensor) did not yield any results when searching via Nexis Uni, and is thus excluded from that section in Figure . To see how these values are related to the patent publishing activity, a comparison is made and displayed in Figure 4.18.

The data from citation- and news-based indicators and comparison with patent publishing activity shows that although the rise of citation index precedes the other two indicators displayed as was expected, occurrences in news articles start to grow before the number of granted patents, contrary to what was expected. In the study by [69], three technologies, biodiesels, laser cladding and blue light-emitting diodes (LEDs) were examined. Interestingly, the first of the three was the only technology of these in which the news indicator started growing before the patent indicator, the same way that the case is with event cameras.

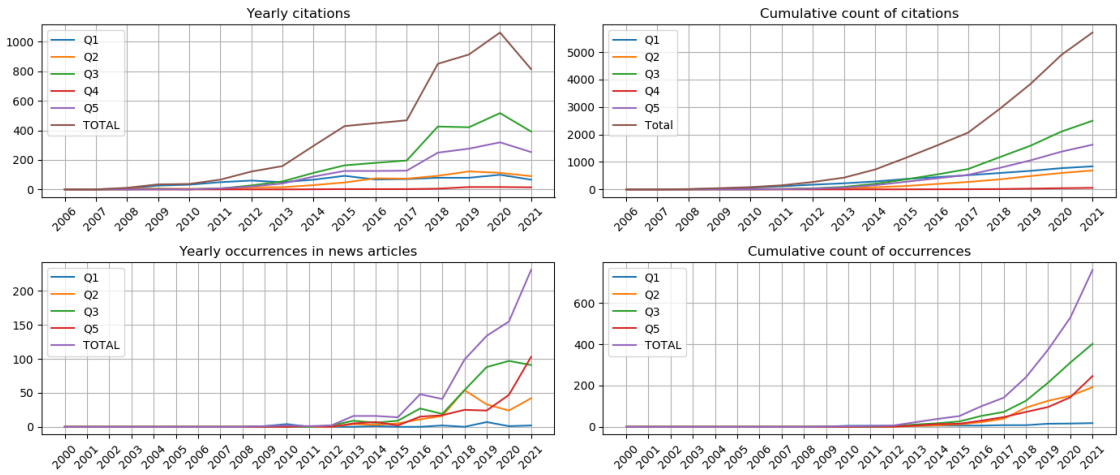

Figure 4.17. *Citation index and occurrences in news articles of queries presented in Table 4.1*

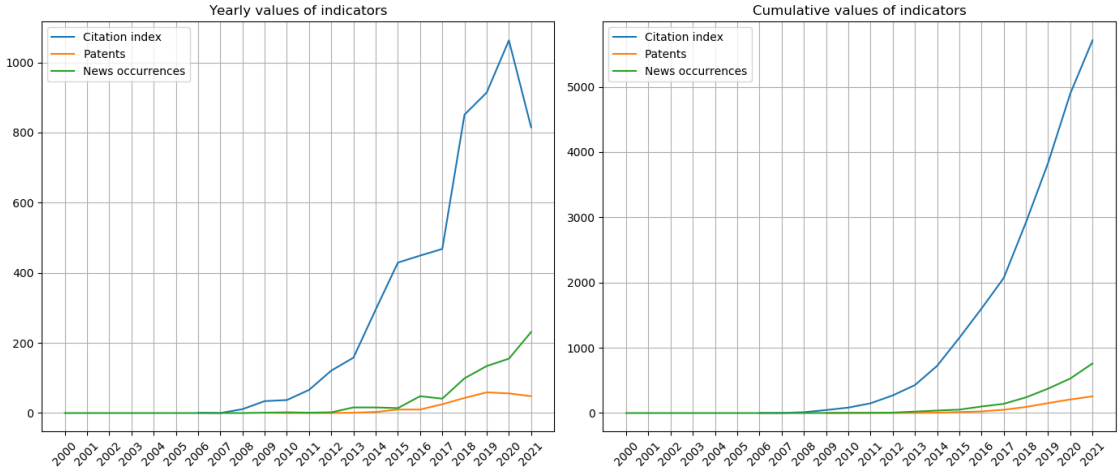

Figure 4.18. *Citation index and occurrences in news articles of queries presented in Table 4.1 compared to patent publishing activity related to the same queries*

4.4 Applications

In addition to granted patents, USPTO database allows users to search for applications under examination. As seen in Figure 4.16, granting a patent can take several years from the date when the patent application is submitted. Searching with the queries that were presented in Table 4.1 from the application database instead of the granted patents database can give some hint on what should be expected from the patents that will be granted in the future. Conducting the same queries for application data yields 367 results, which is 43.4 % increase compared to corresponding granted patents.

Because the motivation is to seek the potential of event camera technology to expand into different industries, especially mobile imaging, it is again important to look at who is applying for the patents related to event camera technology. In Figure 4.19, applicants by

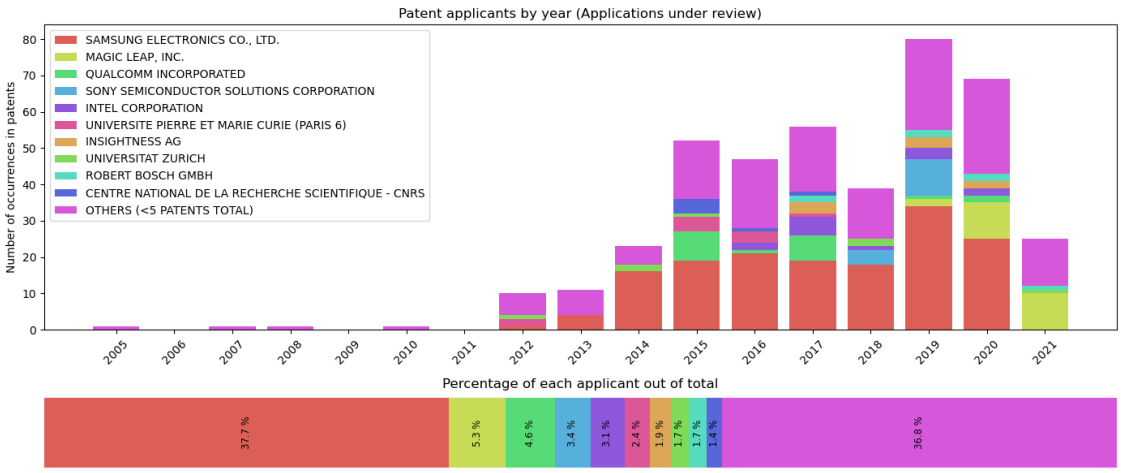

Figure 4.19. *Yearly applicants for non-grated patent applications for event camera technology*

year for non-granted patent applicants related to event camera technology is shown.

Applicant data over the years presented in Figure 4.7 show that Samsung is the top applicant as was the case with granted patents, and also another mobile imaging enterprise (Sony) has submitted dozens of event camera -related patents in recent years. However, examining the contents of the patents by these companies do not suggest that the patents are especially designed for mobile phone imaging applications. R&D efforts on event camera utilization in VR & AR setting should also be noted, as we can see a Magic Leap Inc., an enterprise focusing on those technologies [76] as a major player among the applicants in recent years.

4.5 Summary

Several methods using different metrics were utilized in order to determine which TLC stage the event camera technology is currently on. Methods and their corresponding results are presented in Table 4.6. The analysis shows that the event camera TLC stage is either at late emergence or early growth stage, more likely the latter, and that there are several years of continuous development ahead.

The possible uncertainty of the metrics used should not be dismissed. Especially the time period between 2010 and 2012 when there was no patents granted might cause some misleading results when examining the patent data, since after those years significant changes occur among all indicators when the values rise from zero.

Although event-based vision sensors do have many advantages over RGB sensors, from the perspective of manufacturing they require similar resources. In order to provide a sufficient incentive for existing manufacturers of imaging sensors to allocate some of their currently available or future production capacity towards event sensor production, man-

Method	Stage determined by method
Examining S-curve	Late emergence or growth
Forecast model	Growth
Entropy	Growth
Other indicators	Not conclusive

Table 4.6. *Summary of TLC determination*

ufacturing these sensors should provide them more profits than their current production is generating. The demand for smartphones have been increasing steadily in the past decade and at the same time the number of camera modules in them have been increasing, as shown in Chapter 1, meaning that the manufacturers of imaging sensors that are used in contemporary mobile devices see promising future in the demand of their products, and this suggests that putting resources in the event sensor manufacturing might be too risky on their part. However, some production capacity for event sensors has already been created, since there are several vendors from whom it is possible to purchase the sensors online.

It remains to be seen if enterprises are willing to take the risk and be the early adopters and integrate event sensors in their products. The cost-benefit ratio might still be too high due to the scarce manufacturing capacities. However, in addition to smartphones, forecasts regarding the markets of technologies on which event sensors have the most potential use cases, such as autonomous vehicles and unmanned aerial vehicles show that the market size of those technologies is expected to increase significantly in the coming years [77] [78]. Additionally, the market of smart home security devices, of which the only consumer product with integrated event sensors was part of, is expected to double between 2020 and 2025 as well [79].

In 2020, as the SARS-CoV-2 pandemic shocked the world economy, global computer chip production saw unprecedented demand that it was not able to answer to. As the chip shortage has halted manufacturing of many industries, such as automotive, which is one potential for future event camera markets, it is still unclear how much this will delay the adaptation of new technologies. On the contrary, it would seem feasible to assume that enterprises that face challenges in production due to shortages will use the time on their hands on focusing in R&D activities. If event cameras fail to break through into the consumer products, they still have potential to be used in large-scale industrial applications such as in machine vision for robots, or in aerospace and military applications.

Although the patent data analysis offers lots of details considering event camera development in general, no clear conclusions were able to made from the basis of it for the possible implementations in mobile phones. It was noted that several mobile phone man-

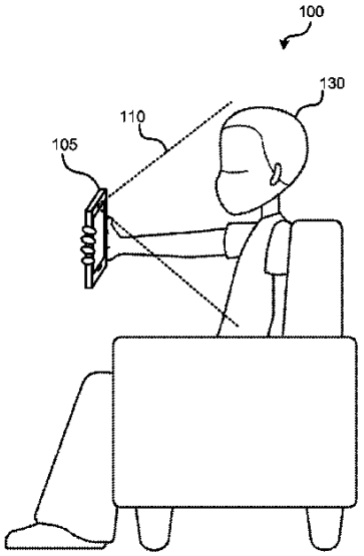

Figure 4.20. *Illustration included in U.S. patent 9554100 [80] depicting a scenario where event sensor can used for face detection and tracking in mobile phone*

ufacturers engage in R&D activities related to event cameras, but the companies in question also research and produce a wide variety of other products in other areas. It should also be noted that some mobile phone manufacturers might not engage in the R&D, but nevertheless implement some features patented by someone else via patent licensing.

Keyword analysis showed that terms related to mobile phones occurred frequently in event camera patent documents, but closer look at the documents revealed that most of the patents are more general in nature and are not especially designated for mobile phones, but that mobile phones are mentioned only in the documents as an example of a device on which the invention could be used in. However, some patents such as [80] clearly depict a concept where event sensors could be utilized in mobile devices, as seen in Figure 4.20. Notably the patent where the illustration in Figure 4.20 is included in, is also included in Table 4.2, where most cited patents in all event camera patents are listed, and is thus considered to be among the most commercially valuable ones. Examining the patents that have cited [80] show that referring patents also have asserted claims on the purpose of utilizing event sensor for iris tracking and iris-based authentication.

It can be argued that the integration of the event sensors into mobile phones is not something that could be patented as itself, or that it is something that the enterprises that considering it would not want to reveal to the public by applying a patent. Debackere et al. [81] have pointed out that in many cases, the decision for not applying for patent protection for an invention is made due to several reasons, which include that some inventions are not patentable, that sometimes acting in secrecy and launching products quickly might result in more success from the business point of view, that the costs of patent

application process are considered too high, or that it is too easy to create competitive innovations that go around the patent. Especially the first, second and last of these reasons can be considered as relevant points from the perspective of mobile phone manufacturing enterprises, when considering whether to apply patents for event camera integration into mobile devices.

5. CONCLUSION

It was demonstrated that event cameras could be used for LSD2-based handshake deblurring applications, but the method does not provide ideal results, since the dynamic range decreases when the data is processed by the network, and some artefacts are present in the outputs. Some possible solutions for the dynamic range problem were found in literature but not tested in this thesis. The method for generating data sets for deblurring purposes by using random motion trajectories and moving window was proven to be realistic enough, since the network trained on that data also provided similar results when the blur was generated from data on which the movement is based on authentic, natural movements rather than artificially generated ones. Overall, the results suggest that event data seems to contain the necessary information for the elimination of handshake blur, making it one possible use case for event cameras in mobile phone among others. However, further research towards finding the best methodologies of utilizing the event data in handshake blur elimination should be conducted.

Event camera technology's TLC was examined and assessments of the possible future outcomes were made. By examining several different indicators derived from patent data, it was concluded that event camera technology is currently most likely at the growth stage, which is considered as the second stage of the TLC. Forecasts based on logistic S-curve model predict that event camera technology will grow for several more years, after it reaches the saturation stage of the life cycle around the year 2025, when the growth starts to slow down. Accuracy of the forecasts models was tested with three technologies that have already reached the saturation stage of their life cycles. Those tests proved the forecast models to be useful tools. Nevertheless the results from forecasts should be considered with caution, especially if the amount of real observations that are used to make the forecast is little, or if the technology in question has not started growing rapidly.

Although several enterprises which have strong interests in mobile phone imaging sector have applied and been granted patents related to event camera technology, the contents of patents and applications by those enterprises mostly imply that the patents are for more generic usage but could possibly be used in mobile phones. This may be due to the fact that patented algorithms and applications that could be used within mobile phones with integrated event sensors could be used elsewhere as well. However, one patent among the most valuable ones measured by the number of forward citations does

depict a scenario where event sensor is used in a mobile phone. Areas other than mobile phone imaging stand out when examining the patent contents. Technologies such as autonomous vehicles and virtual or augmented reality are at the core of multiple patents from several different applicants. This suggests that event camera technologies might be integrated in products in these technology categories in the coming years.

Tools were developed to be used in automatized patent document retrieval from the USPTO patent and application databases and for parsing important information from the retrieved documents. These tools could be easily utilized in the future when inspecting other technologies of interest by only changing the queries that were used to search for the data. Building new tools for TLC and other types of patent-based analysis on top of the environment that was created could also be done.

REFERENCES

[1] Statista. *Number of smartphones sold to end users worldwide from 2007 to 2021*. 2021. URL: https://www.statista.com/statistics/263437/global-smartphone-sales-to-end-users-since-2007 (visited on 09/28/2021).

[2] Statista. *Smartphone sales forecasts in the United States from 2005 to 2021*. 2021. URL: https://www.statista.com/statistics/191985/sales-of-smartphones-in-the-us-since-2005/ (visited on 09/28/2021).

[3] Statista. *Size of the camera module market in North America by application from 2012 to 2022*. 2021. URL: https://www.statista.com/statistics/691640/camera-module-market-size-by-application-in-north-america/ (visited on 09/28/2021).

[4] Statista. *Average price of smartphones in the United States from 2014 to 2024, by segment*. 2021. URL: https://www.statista.com/statistics/619830/smartphone-average-price-in-the-us/ (visited on 09/28/2021).

[5] Statista. *Tablet market in the U.S.* 2020. URL: https://www.statista.com/study/10519/us-tablet-pc-market-statista-dossier/.

[6] Gallego, G., Delbruck, T., Orchard, G. M., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A., Conradt, J., Daniilidis, K. and Scaramuzza, D. Event-based Vision: A Survey. eng. *IEEE transactions on pattern analysis and machine intelligence* PP (2020), pp. 1–1. ISSN: 0162-8828.

[7] Posch, C. Bio-inspired vision. eng. *Journal of instrumentation* 7.1 (2012), pp. C01054–C01054. ISSN: 1748-0221.

[8] Lichtsteiner, P., Posch, C. and Delbruck, T. A 128×128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor. *IEEE Journal of Solid-State Circuits* 43.2 (2008), pp. 566–576. DOI: 10.1109/JSSC.2007.914337.

[9] Brandli, C., Berner, R., Yang, M., Liu, S.-C. and Delbruck, T. A 240×180 130 dB 3 μs Latency Global Shutter Spatiotemporal Vision Sensor. *IEEE Journal of Solid-State Circuits* 49.10 (2014), pp. 2333–2341. DOI: 10.1109/JSSC.2014.2342715.

[10] Posch, C., Matolin, D. and Wohlgenannt, R. A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS. *IEEE Journal of Solid-State Circuits* 46.1 (2011), pp. 259–275. DOI: 10.1109/JSSC.2010.2085952.

[11] AlpsenTek. *Products*. 2021. URL: https://www.alpsentek.com/product/.

[12] Son, B., Suh, Y., Kim, S., Jung, H., Kim, J.-S., Shin, C., Park, K., Lee, K., Park, J., Woo, J., Roh, Y., Lee, H., Wang, Y., Ovsiannikov, I. and Ryu, H. A 640×480 dynamic

vision sensor with a 9μm pixel and 300Meps address-event representation. eng. *Digest of Technical Papers - IEEE International Solid-State Circuits Conference*. Vol. 60. 2017, pp. 66–67. ISBN: 9781509037575.

[13] Finateu, T., Niwa, A., Matolin, D., Tsuchimoto, K., Mascheroni, A., Reynaud, E., Mostafalu, P., Brady, F., Chotard, L., LeGoff, F., Takahashi, H., Wakabayashi, H., Oike, Y. and Posch, C. 5.10 A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with 4.86µm Pixels, 1.066GEPS Readout, Programmable Event-Rate Controller and Compressive Data-Formatting Pipeline. eng. *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*. IEEE, 2020, pp. 112–114. ISBN: 1728132053.

[14] Suh, Y., Choi, S., Ito, M., Kim, J., Lee, Y., Seo, J., Jung, H., Yeo, D.-H., Namgung, S., Bong, J., Yoo, S., Shin, S.-H., Kwon, D., Kang, P., Kim, S., Na, H., Hwang, K., Shin, C., Kim, J.-S., Park, P. K. J., Kim, J., Ryu, H. and Park, Y. A 1280×960 Dynamic Vision Sensor with a 4.95-μm Pixel Pitch and Motion Artifact Minimization. eng. *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2020, pp. 1–5. ISBN: 9781728133201.

[15] Samsung. *Samsung*. 2021. URL: https://www.samsung.com/ (visited on 09/28/2021).

[16] Bolten, T., Pohle-Fröhlich, R. and Tönnies, K. D. DVS-OUTLAB: A Neuromorphic Event-Based Long Time Monitoring Dataset for Real-World Outdoor Scenarios. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2021, pp. 1348–1357. DOI: 10.1109/CVPRW53098.2021.00149.

[17] Sokolova, A. and Konushin, A. Human identification by gait from event-based camera. eng. *2019 16th International Conference on Machine Vision Applications (MVA)*. MVA Organization, 2019, pp. 1–6. ISBN: 4901122185.

[18] Sarmadi, H., Munoz-Salinas, R., Olivares-Mendez, M. A. and Medina-Carnicer, R. Detection of Binary Square Fiducial Markers Using an Event Camera. eng. *IEEE access* 9 (2021), pp. 27813–27826. ISSN: 2169-3536.

[19] Falanga, D., Kleber, K. and Scaramuzza, D. Dynamic obstacle avoidance for quadrotors with event cameras. eng. *Science robotics* 5.40 (2020), eaaz9712–. ISSN: 2470-9476.

[20] Aeronautics, N. and Administration, S. *NASA's Ingenuity Mars Helicopter Succeeds in Historic First Flight*. 2021. URL: https://www.nasa.gov/press-release/nasa-s-ingenuity-mars-helicopter-succeeds-in-historic-first-flight (visited on 09/28/2021).

[21] Muglikar, M., Moeys, D. P. and Scaramuzza, D. Event Guided Depth Sensing. eng. (2021).

[22] Statista. *Smartphone camera technology penetration rate worldwide from 2016 to 2019*. 2018. URL: https://www.statista.com/statistics/947043/worldwide-penetration-rate-smartphone-camera-technology/ (visited on 10/13/2021).

[23] Statista. *Number of digital single-lens reflex cameras sold in Norway from 2009 to 2020*. 2021. URL: https://www.statista.com/statistics/947043/640256/number-of-dslr-cameras-sold-in-norway/ (visited on 10/13/2021).

[24] Statista. *Sales volume of digital single-lens reflex (SLR) cameras and compact system cameras in Germany from 2006 to 2020*. 2021. URL: https://www.statista.com/statistics/771786/digital-reflex-system-cameras-sales-volume-germany/ (visited on 10/13/2021).

[25] Rebecq, H., Ranftl, R., Koltun, V. and Scaramuzza, D. High Speed and High Dynamic Range Video with an Event Camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.6 (2021), pp. 1964–1980. DOI: 10.1109/TPAMI.2019.2963386.

[26] Ryan, C., O'Sullivan, B., Elrasad, A., Cahill, A., Lemley, J., Kielty, P., Posch, C. and Perot, E. Real-time face & eye tracking and blink detection using event cameras. eng. *Neural networks* 141 (2021), pp. 87–97. ISSN: 0893-6080.

[27] Lenz, G., Ieng, S.-H. and Benosman, R. Event-Based Face Detection and Tracking Using the Dynamics of Eye Blinks. eng. *Frontiers in neuroscience* 14 (2020), pp. 587–587. ISSN: 1662-453X.

[28] Chen, J., Meng, J., Wang, X. and Yuan, J. Dynamic Graph CNN for Event-Camera Based Gesture Recognition. eng. *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2020, pp. 1–5. ISBN: 9781728133201.

[29] Pan, L., Hartley, R., Scheerlinck, C., Liu, M., Yu, X. and Dai, Y. High Frame Rate Video Reconstruction based on an Event Camera. eng. *IEEE transactions on pattern analysis and machine intelligence* PP (2020), pp. 1–1. ISSN: 0162-8828.

[30] Kim, H., Leutenegger, S. and Davison, A. J. Real-Time 3D Reconstruction and 6-DoF Tracking with an Event Camera. *Computer Vision – ECCV 2016*. Ed. by B. Leibe, J. Matas, N. Sebe and M. Welling. Cham: Springer International Publishing, 2016, p. 351. ISBN: 978-3-319-46466-4.

[31] Rebecq, H., Gehrig, D. and Scaramuzza, D. ESIM: an Open Event Camera Simulator. *Conf. on Robotics Learning (CoRL)* (Oct. 2018).

[32] Research, M. AirSim. 2020. URL: https://microsoft.github.io/AirSim/event_sim/ (visited on 10/28/2021).

[33] Tournemire, P. de, Nitti, D., Perot, E., Migliore, D. and Sironi, A. *A Large Scale Event-based Detection Dataset for Automotive*. 2020. arXiv: 2001.08499 [cs.CV].

[34] Miao, S., Chen, G., Ning, X., Zi, Y., Ren, K., Bing, Z. and Knoll, A. Neuromorphic Vision Datasets for Pedestrian Detection, Action Recognition, and Fall Detection. *Frontiers in Neurorobotics* 13 (2019), p. 38. ISSN: 1662-5218. DOI: 10.3389/fnbot.2019.00038. URL: https://www.frontiersin.org/article/10.3389/fnbot.2019.00038.

[35] Delmerico, J., Cieslewski, T., Rebecq, H., Faessler, M. and Scaramuzza, D. Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset. *IEEE Int. Conf. Robot. Autom. (ICRA)*. 2019.

[36] Šindelář, O., Šroubek, F. and Milanfar, P. A smartphone application for removing handshake blur and compensating rolling shutter. *2014 IEEE International Conference on Image Processing (ICIP)*. 2014, pp. 2160–2162. DOI: 10.1109/ICIP.2014.7025433.

[37] Yang, F.-W., Lin, H. J. and Chuang, H. Image deblurring. eng. *2017 IEEE SmartWorld, Ubiquitous Intelligence& Computing, Advanced& Trusted Computed, Scalable Computing& Communications, Cloud& Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*. IEEE, 2017, pp. 1–4. ISBN: 1538604353.

[38] Zhang, L., Zhang, H., Chen, J. and Wang, L. Hybrid Deblur Net: Deep Non-Uniform Deblurring With Event Camera. *IEEE Access* 8 (2020), pp. 148075–148083. DOI: 10.1109/ACCESS.2020.3015759.

[39] Mustaniemi, J., Kannala, J., Matas, J., Särkkä, S. and Heikkilä, J. LSD2 Joint Denoising and Deblurring of Short and Long Exposure Images with CNNs. *The 31st British Machine Vision Virtual Conference (BMVC)*. Sept. 2020.

[40] Ronneberger, O., Fischer, P. and Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. eng. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Vol. 9351. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 9783319245737.

[41] Huiskes, M. J. and Lew, M. S. The MIR Flickr Retrieval Evaluation. *MIR '08: Proceedings of the 2008 ACM International Conference on Multimedia Information Retrieval*. Vancouver, Canada: ACM, 2008.

[42] Boracchi, G. and Foi, A. Modeling the Performance of Image Restoration From Motion Blur. *IEEE Transactions on Image Processing* 21.8 (2012), pp. 3502–3517. DOI: 10.1109/TIP.2012.2192126.

[43] Paalanen, P. and Kämäräinen, J. *Camera Movement Emulator*. 2021. URL: https://github.com/kamarain/camera_movement_emulator/.

[44] Wang, Z., Bovik, A., Sheikh, H. and Simoncelli, E. Image quality assessment: from error visibility to structural similarity. eng. *IEEE transactions on image processing* 13.4 (2004), pp. 600–612. ISSN: 1057-7149.

[45] Nilsson, J. and Akenine-Möller, T. *Understanding SSIM*. 2020. arXiv: 2006.13846 [eess.IV].

[46] Nah, S., Kim, T. H. and Lee, K. M. Deep Multi-Scale Convolutional Neural Network for Dynamic Scene Deblurring. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 2017.

[47] Marnerides, D., Bashford-Rogers, T. and Debattista, K. Spectrally Consistent UNet for High Fidelity Image Transformations. eng. (2020).

[48] Taylor, M. and Taylor, A. The technology life cycle: Conceptualization and managerial implications. eng. *International journal of production economics* 140.1 (2012), pp. 541–553. ISSN: 0925-5273.

[49] Ernst, H. Patent information for strategic technology management. *World Patent Information* 25.3 (2003), pp. 233–242. ISSN: 0172-2190. DOI: https://doi.org/10.1016/S0172-2190(03)00077-2. URL: https://www.sciencedirect.com/science/article/pii/S0172219003000772.

[50] The use of S curves in technology forecasting and its application on 3D tv technology. *Proceedings of World Academy of Science, Engineering and Technology* 6.11 (2012), pp. 1283–1287. ISSN: 2010-376X.

[51] Anderson, P. and Tushman, M. L. Technological discontinuities and dominant designs - a cyclical model of technological change. English. *Administrative Science Quarterly* 35.4 (Dec. 1990). Date revised - 2013-06-12; Last updated - 2013-09-16; SubjectsTermNotLitGenreText - 3245 971; 12616 12622, pp. 604–633. URL: https://libproxy.tuni.fi/login?url=https://www-proquest-com.libproxy.tuni.fi/scholarly-journals/technological-discontinuities-dominant-designs/docview/839232640/se-2?accountid=14242.

[52] Altuntas, S., Dereli, T. and Kusiak, A. Forecasting technology success based on patent data. eng. *Technological forecasting & social change* 96 (2015), pp. 202–214. ISSN: 0040-1625.

[53] Lin, D., Liu, W., Guo, Y. and Meyer, M. Using technological entropy to identify technology life cycle. *Journal of Informetrics* 15.2 (2021), p. 101137. ISSN: 1751-1577. DOI: https://doi.org/10.1016/j.joi.2021.101137. URL: https://www.sciencedirect.com/science/article/pii/S1751157721000080.

[54] Boundless. *The Technology Life Cycle*. 2021. URL: http://kolibri.teacherinabox.org.au/modules/en-boundless/www.boundless.com/management/textbooks/boundless-management-textbook/organizational-culture-and-innovation-4/technology-and-innovation-37/the-technology-life-cycle-202-3486/index.html (visited on 09/28/2021).

[55] USPTO. *USPTO full-text database*. 2021. URL: https://patft.uspto.gov/netahtml/PTO/index.html (visited on 08/10/2021).

[56] EPO. *Espacenet database*. 2021. URL: https://worldwide.espacenet.com/ (visited on 08/10/2021).

[57] Criscuolo, P. The 'home advantage' effect and patent families. A comparison of OECD triadic patents, the USPTO and the EPO. eng. *Scientometrics* 66.1 (2006), pp. 23–41. ISSN: 0138-9130.

[58] Statista. *Global market share of the information and communication technology (ICT) market from 2013 to 2021, by selected country*. 2021. URL: https://www.statista.com/statistics/263801/global-market-share-held-by-selected-countries-in-the-ict-market (visited on 09/28/2021).

[59] Muthukadan, B. *Selenium Documentation*. 2021. URL: https://selenium-python.readthedocs.io/ (visited on 08/10/2021).

[60] Richardson, L. *BeautifulSoup 4 Documentation*. 2021. URL: https://beautiful-soup-4.readthedocs.io/en/latest/ (visited on 08/10/2021).

[61] Patent indicators for the technology life cycle development. eng. *Research policy*. Research Policy 36.3 (2007), pp. 387–398. ISSN: 0048-7333.

[62] USPTO. *USPTO Stopwords*. 2021. URL: https://patft.uspto.gov/netahtml/PTO/help/stopword.htm (visited on 08/10/2021).

[63] USPTO. *A-Z definitions of intellectual property terminology.* 2020. URL: https://www.uspto.gov/learning-and-resources/glossary (visited on 09/28/2021).

[64] Hasan, F. *Harvard Business School - Baker Library - Fast Answers*. 2020. URL: https://asklib.library.hbs.edu/faq/267606 (visited on 09/28/2021).

[65] Henderson, R., Jaffe, A. B. and Trajtenberg, M. Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988. eng. *The review of economics and statistics* 80.1 (1998), pp. 119–127. ISSN: 0034-6535.

[66] Organization, W. I. P. *International Patent Classification (IPC)*. 2021.

[67] Statista. *Statista dossier about Qualcomm*. 2021.

[68] Thomas, J. J., Karagoz, P., Ahamed, B. B. and Vasant, P. *11.9 TF-IDF*. IGI Global, 2020, p. 172. ISBN: 978-1-79981-192-3. URL: https://app.knovel.com/hotlink/khtml/id:kt012675A2/deep-learning-techniques/tf-idf.

[69] Järvenpää, H., Mäkinen, S. and Seppänen, M. *Patent and publishing activity sequence over a technology's life cycle*. eng. Tampere University, 2011.

[70] Huang, L. and Wang, N. Status and Prospects of Technology Diffusion Research based on Patent Information. eng. *Proceedings of the Sixth International Conference on Management Science and Engineering Management*. Vol. 185. Lecture Notes in Electrical Engineering. London: Springer London, 2012, pp. 167–180. ISBN: 9781447145998.

[71] Statista. *Display technology - Statistics & Facts*. 2020.

[72] Statista. *Flash memory market revenues worldwide from 2013 to 2021*. 2017. URL: https://www.statista.com/statistics/553556/worldwide-flash-memory-market-size/ (visited on 09/28/2021).

[73] Liu, C.-Y. and Wang, J.-C. Forecasting the development of the biped robot walking technique in Japan through S-curve model analysis. eng. *Scientometrics* 82.1 (2010), pp. 21–36. ISSN: 0138-9130.

[74] Chen, Y. H., Chen, C. Y. and Lee, S. C. Technology forecasting and patent strategy of hydrogen energy and fuel cell technologies. *International Journal of Hydrogen Energy* 36.12 (2011), pp. 6957–6969. ISSN: 03603199. DOI: 10.1016/j.ijhydene.2011.03.063. URL: http://dx.doi.org/10.1016/j.ijhydene.2011.03.063.

[75] LexisNexis. *Nexis Uni: Academic Research Reimagined for Digital Natives*. 2021. URL: https://www.lexisnexis.com/en-us/professional/academic/nexis-uni.page (visited on 10/08/2021).

[76] Incorporated, M. L. *Magic Leap Incorporated*. 2021. URL: https://www.magicleap.com/en-us (visited on 10/08/2021).

[77] Statista. *Projected sales of autonomous vehicles worldwide from 2019 to 2030*. 2021. URL: https://www.statista.com/statistics/1230733/projected-sales-autonomous-vehicles-worldwide/ (visited on 09/28/2021).

[78] Statista. *Drone market revenue worldwide from 2019 to 2025*. 2020. URL: https://www.statista.com/statistics/1200348/drone-market-revenue-worldwide/ (visited on 09/28/2021).

[79] Statista. *Smart home security market revenue worldwide from 2018 to 2025*. 2021. URL: https://www.statista.com/statistics/1056057/worldwide-smart-home-security-market-value/ (visited on 09/28/2021).

[80] Gousev, E. P., Govil, A., Maitan, J., Rasquinha, N. and Rangan, V. *U.S. Patent 9,554,100: Low-power always-on face detection, tracking, recognition and/or analysis using events-based vision sensor*. 2017.

[81] Debackere, K., Verbeek, A., Luwel, M. and Zimmermann, E. Measuring progress and evolution in science and technology - II: The multiple uses of technometric indicators. eng. *International journal of management reviews : IJMR* 4.3 (2002), pp. 213–231. ISSN: 1460-8545.