COMP0073 – MSc Computer Science Project

An Investigation Into Possible Use Cases of the Azure Kinect DK Supported by an Analysis of the IoT Industry and key Players

Research Deliverable to the Industry Partner Avanade

Constantin Ulbrich
Student Number: 15010034

Internal Supervisor: Prof. Dr. Yuzuko Nakamura
External Supervisor: Fergus Kidd
Degree: MSc Computer Science
Submission Date: September 11, 2020

Department of Computer Science
University College London

Table of Content

List of Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
AI	Artificial Intelligence
CAGR	Compounded Annual Growth Rate
CRE	Commercial Real Estate
EBCA	Equal Baseline Camera Array
fps	Frames per second
HADL	Human Activities of Daily Living
HCI	Human-Computer Interaction
HRI	Human-Robot Interaction
IIoT	Industrial Internet of Things
IMU	Inertial Measurement Unit
IoT	Internet of Things
IR	Infrared
IT	Information Technology
KPI	Key Performance Index
LiDAR	Light Detection and Ranging
RA	Rheumatoid Arthritis
RGBD camera	Red Green Blue Depth camera
SaaS	Software as a Service
SDK	Software Developer Kit
ToF	Time-of-Flight

List of Figures

List of Graphs

List of Tables

1 Introduction

This research investigation is being conducted as part of the Smart Vision Project. The project aimed to propose and validate a new smart space service, using passive observer technology, for the Modern Workplace Practice of the technology consultancy Avanade. Specifically, the client was interested in learning about potential uses of the just released Azure Kinect DK for such a smart space application. To explore the potential of a technology for a novel service, the industry project was divided into a research and proof of concept deliverable.

1.1 Task/Problem Statement

This investigation is the first deliverable of the Smart Vision Project. It focuses on the Microsoft Azure Kinect DK sensor camera. The positioning of the Kinect DK against competitive products and potential use cases for the sensor camera are explored. Moreover, insights about key players in the Smart Space/IoT industry are collected and analysed. To achieve those targets, the paper consists of several different types of analysis in order to capture various point of views on the discussed topics. To ensure an alignment across the various analysis, the research goals were rephrased into three research questions:

1. What kind of uses cases fit to the functionality profile of the Azure Kinect DK, as a passive observer technology, and could be of relevance to Avanade's Emerging Technology Team?
2. How does the Kinect DK's functionalities compare to sensors with similar functionalities?
3. How does the current market environment of the IoT industry look like?

Over and above, the first deliverable of the Smart Vision Project, this investigation also contributes key inputs for the second deliverable. For instance, the competitive product mapping ensures that the Azure Kinect is a strong performing visual sensor and the technical analysis of the depth sensing technologies allows to evaluate a use case fit. Most importantly, the findings of this investigation determine the type of proof of concept application, which is built for the second project deliverable.

1.2 Scope of the Investigation

The overall research scope of this investigation is rather broad. There is no geographic scope, and the time scope of the investigation was set to be as recent as possible. However, precise boundaries were determined by the types of analysis conducted in this investigation. The research questions should be fully addressed by the findings of a competitive product mapping, a systematic literature review, as well as an industry and competitor analysis. Due to the breadth of certain investigated fields, the scope was narrowed within each of the analysis respectively. For example, the scope was confined as a result of clearly defined inclusion and exclusion criteria, or by defining an analysis sample.

1.2 Course of Investigation

The investigation continues with the positioning of the Azure Kinect DK against competitor sensors in form of a competitive product mapping. Next, the identified depth sensor technologies from the competitive product mapping are discussed on a technical level. Subsequently, a systematic literature review is conducted to investigate current trends in academia for computer vision technology. Then, the research questions are tackled from an industry perspective in form of an industry analysis of the smart spaces/IoT market and an analysis of the key players and their product offerings. In the next step, the findings of the different types of analysis are synthesized and presented in form of a use case fit description, which is then used to propose key use cases for the Azure Kinect DK. Lastly, the legal and ethical implications are highlighted, as well as limitations of the investigation are pointed out.

2 Technical Analysis of the Azure Kinect DK and Competitive Devices

As part of the research component's goals, it is of Avanade's interest to evaluate the Azure Kinect DK sensor against competitor sensors capable of achieving similar tasks. To achieve this requirement, a detailed market analysis was conducted. The Kinect DK sensor combines several sensors in one device. Consequently, it offers many different applications, touching the sensor categories Image, 3D/Depth, Motion, and Microphone. To refine the product research, the investigation focuses on devices with a similar set of integrated sensors. Specific weight was given to Image and 3D/Depth sensors, as Avanade tasked to explicitly explore use cases with the Kinect DK involving smart vision in terms of human and object detection, as well as motion tracking. The scope was yielding too many results to be reviewed; therefore, the sample set was further narrowed down by excluding sensors only designed for robotics, setting a maximum price of $500, and technical information about the sensor had to be freely available. Further guidance was provided by ROS-Industrial's (2020) ongoing comparison of 3D/Depth sensors. This led to a sample result of eight popular Image and 3D/Depth sensors, which are compared in the table below.

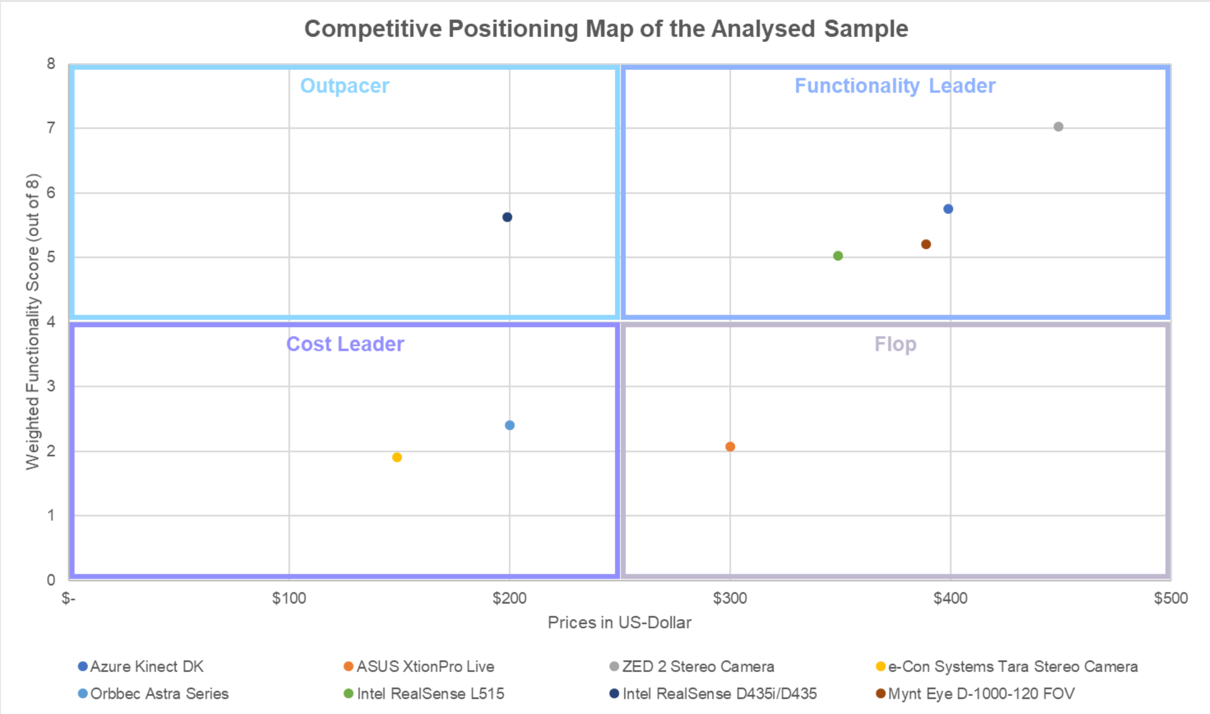
Table 1. Sample of competitive products to the Microsoft Azure Kinect DK.

Sensor Name	Azure Kinect DK	ASUS XtionPro Live	ZED 2 Stereo Camera	e-Con Systems Tara Stereo Camera	Orbbec Astra Series	Intel RealSense L515	Intel RealSense D435i/D435	Mynt Eye D-1000-120 FOV
Sensor Image								
RGB Camera	Yes (max. resolution: 4096 x 3072, 12 MP)	Yes (max. Resolution: 1280 x 1024)	Yes (max. Resolution: 2208 x 1242)	Yes	Yes (max. Resolution: 1280 x 720)	Yes (max. resolution: 1920 x 1080, 2 MP)	Yes (max. Resolution: 1920 x 1080)	Yes (max. Resolution: 2560 x 720)
Frame Rate	15-60 fps.	30 fps.	15-100 fps.	n.a.	30 fps.	30-60 fps.	30 fps.	60 fps.
Depth-Sensor	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Type	Time-of-Flight	Structured Light	Embedded Stereo	n.a.	Embedded Stereo	LiDAR	Active IR Stereo	Embedded Stereo
Depth-Range	0.5 - 4.5 m	0.8 - 3.5 m	0.3 - 20 m	0.5 - 3 m	0.6 - 8 m	0.25 - 9 m	0.1 - 10 m	0.5 – 15 m
3D-Resolutions	1024 x 1024	640 x 480	2208 x 1242	752 x 480	640 x 480	1024 x 768	1280 x 720	1280 x 720
Frame Rate	30 fps	30 fps	100 fps.	60 fps.	30 fps	30 fps	90 fps.	60 fps.
Microphone	360-degree 7-microphone array	Yes	No	No	Yes	No	No	No
Accelerometer	Yes	No	Yes	No	No	Yes	Yes	Yes
Gyroscope	Yes	No	Yes	No	No	Yes	Yes	Yes
Other Sensors	No	No	Yes	No	No	No	No	No
SDK	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Physical Dimensions (in mm)	130 x 39 x 126	180 x 35 x 50	175 x 30 x 33	100 x 30 x 35	165 x 30 x 40	61 x 26	90 x 25 x 25	165 x 30 x 30
Weight	440g	540g	166g	80.5g	300g	95g	72g	194g
Connector	USB-C	USB 2.0/3.0	USB 3.0	USB 2.0/3.0	USB 2.0	USB-C	USB-C	USB 3.0
Price	$399	$300	$449	$149	ca. $200	$349	$199/$179	$389
Source	Microsoft, 2020	Asus, (n.a.); ROS-Industrial (2020)	StereoLabs (2020); ROS-Industrial, (2020)	e-con Systems, (2020); ROS-Industrial, (2020)	Orbbec (n.a.); ROS-Industrial, (2020)	Intel (2020)	Intel (2019A)	Mynt Eye (n.a.); ROS-Industrial, (2020)

Some key observations can be drawn from the table. The integrated depth-sensors deploy four different technologies to generate depth-maps. The most popular method among the sample is stereo vision, either embedded in the RGB cameras or by using Infrared sensors. The other methods to measure depth are Time-of-Flight, Structured Light, and LiDAR. Those technologies will be discussed in later sections. The depth-range of the devices can be associated to three groups. The first group operates in close distance with a maximum depth of about 5 meters, the second covers up to 10 meters, and two devices go up to 15 and 20 meters. There is greater consistency among the frame rate for the RGB camera and depth sensor. The current state of technology allows 30 fps for an average operating mode across the majority of devices in this sample. Almost all of the expensive devices allow to manipulate the frame rate with an inverse relationship to the image resolution. A distinction factor among the products appears to be additional sensors. Only three devices have microphones. Accelerometers and gyroscopes are more common, but do not seem to be industry practice yet. Only one other device integrated three additional sensors. Judging the value gains by integrating additional sensors does not offer a clear answer. It depends on the use cases of the devices. All device manufactures offer an SDK. Key differentiator is not the SDK itself, but the ease of combining different software services with it. The large technology companies in the sample have a clear lead in this area. Microsoft and Intel are providing their Cognitive Services and OpenVINO toolkits respectively to streamline the development of machine learning vision and computer vision applications. These key observations help to paint a general picture of the solutions in the Depth/3D camera market. However, to gain in-depth market research insights a competitive positioning map was constructed (D'Aveni, 2007).

This competitive positioning map plots a product's functionality score against its market price. The functionality score was derived from the technical criteria captured in the table. For each criterion the devices were ranked. The ranking was then mapped to a specific score. Finally, a weighted scoring system was used to obtain the final functionality score, which has a value range between 0 (worst) and 8 (best). A detailed explanation for the calculation of the functionality score can be found in appendix A. The graph below depicts the competitive positioning map (D'Aveni, 2007).

The competitive positioning is partitioned into four categories: outpacer, functionality leader, cost leader, and flops. The cost leaders were the Tara Stereo Camera by e-Con Systems and Orbbec's Astra Series. Both devices have limited functionality but at a very attractive price. The functionality leader group is dominated by the Zed 2 Stereo Camera. It offers the most and best functionality, however at the highest cost. The device is followed by the Azure Kinect DK, the Mynt Eye D, and Intel Real Sense L515. The only outpacer is the RealSense D435i/435, manufactured by Intel. It offers the same level of functionality as the three latter named devices at nearly half the price. The Depth/3D camera sold by Asus is a flop, as it delivers the second worse functionality at almost the price of the functionality leader segment.



Graph 1. Competitive Positioning Map of Azure Kinect DK's competitors.

Overall, the Microsoft Kinect is the second-best sensor device in terms of functionality. It stands out with its seven-array microphone, enabling a 360-degree coverage. No other device can offer this. In combination with its high resolution RGB camera and medium range depth sensors, as well as its gyroscope and accelerometer, the Kinect is a strong allrounder. Such a device excels in applications, which have to combine several sensor functionalities and a medium operational distance. For example, the Kinect would best be applied in a use case requiring computer vision and speech recognition functionalities. However, there are cheaper options available if a use case only requires a particular type of sensor for a successful execution. Furthermore, the outpacer Real Sense D435i/435 is a severe competition to the Kinect, as it offers nearly the same functionalities at half the price.

3 Analysis of Identified Depth/3D Sensor Technologies

The competitive product mapping, conducted in section 2, identified four different types of Depth/3D Sensor technologies among the sample devices: Time-of-Flight, Embedded Stereo, Structured Light, LiDAR, and Active IR Stereo. The following section explores these types of sensor technologies in greater detail from a technical point of view. This should help to gain a better understanding of the application areas of the different types of technologies and support the process to determine technological characteristics for a use case fit. The figure below depicts the identified depth sensing technologies found in the sensor sample from section 2. It also highlights the relations to each other and the underlying base technology.

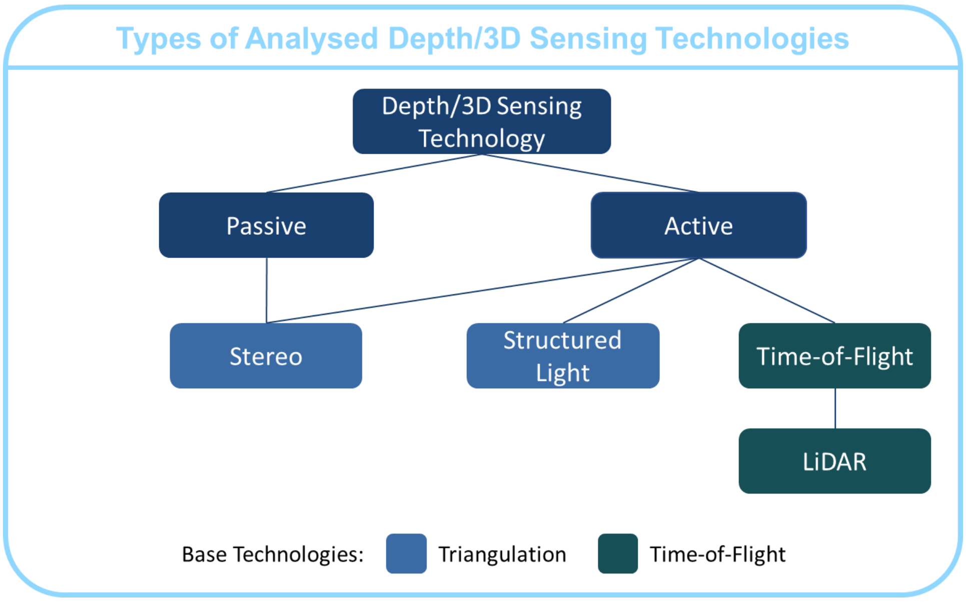


Figure 1. Types of analysed Depth/3D Sensing Technologies adapted from Anderson (2016).

3.1 Depth-/3D-Sensor Technology: Time-of-Flight & LiDAR

Time-of-Flight (ToF) is a depth range imaging technique, making use of the basic properties of light. The ToF technique has two simple but very effective methods to measure depth ranges. Both techniques emit a light ray from a light source. The first method measures the time it takes for the ray to travel back and forth between the sensor system and an object (Kim & Lee, 2013). This enables to calculate the depth, as the speed of light is known and the distance from the emitter to the object equates to the depth: $d = \frac{c \times t}{2}$, where d is the distance in meters, c equates to the speed of light in meters per second, and t represents the time travelled in seconds (Kim & Lee, 2013). The second method determines the depth by calculating the phase shift of the emitted and received light wave, when it is reflected by an object (Kim & Lee, 2013). From the phase shift the distance to the object is determined. By deriving the depth from the phase shift, the sensor device would not require a high-precision clock (Kim & Lee, 2013). This would make the hardware costs even cheaper. Additionally, to establish a depth map of the surrounding the first method's accuracy relies on the light emitter's ability to pulse at high frequencies. The figure below visualizes the two ToF methods.

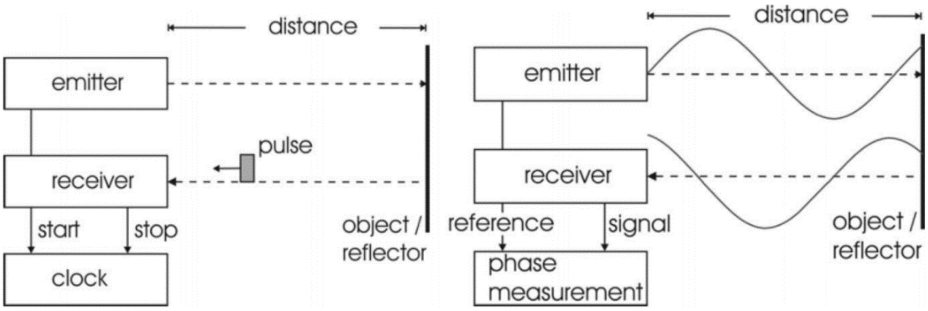


Figure 2. ToF travel time (left) and Phase Shift Method (right) (Kim & Lee, 2013).

The ToF depth range imaging technique can be implemented in various kinds of formats. One major differentiating component is the light emitter. Many systems operate with infrared light (Tillman, 2020). For example, a ToF system, which obtained significant attention due to its extensive use in autonomous driving, is LiDAR. A pulsed infrared laser sweeps the surroundings to establish with the described ToF technique a depth-map (Sharma, 2019).

3.2 Depth-/3D-Sensor Technology: Structured Light

Structured Light depth imaging is derived from the basic concept of triangulation. It consists of a photodetector and a light source (Myers, 2018). The source projects a pattern of known dimensions and geometry onto the surroundings. The projected pattern is distorted by any objects, which it encounters. The photodetector senses the newly formed pattern and algorithms use the light waves' phase shift and deviations of the regular geometry to calculate a depth map of the illuminated surrounding (Myers, 2018). The light source can operate with different types of light. Frequently, Structured Light depth imaging is carried out with infrared light (Myers, 2018), but also high intensity (visible) lasers are used. The figure below depicts the Structured Light technique.

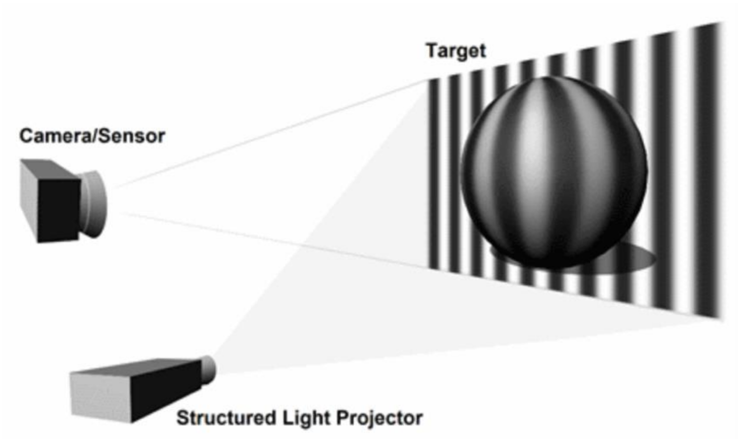


Figure 3. Visualization of the Structured Light Depth Imaging Technique (MoviMed, 2018).

3.3 Depth-/3D-Sensor Technology: Embedded Stereo & Active IR Stereo

Stereo imaging is derived from the technique of stereoscopy, which in turn applies the basic principles of triangulation. On a high-level, this technique combines two 2-D images, taken synchronously off the same surrounding but from different angles, to a 3-D image. This introduces the depth component, or z-axis. The technique of stereoscopy is reconstructed in a Stereo imaging sensor. The device consists of two cameras mounted in parallel. An image is taken by each of the cameras simultaneously of the same scene (Ambrosch, Humenberger, Olufs & Schraml, 2010). Each point in the scene is captured by an active-pixel chip, meaning that each scene point is represented by a pixel in the image (O' Riordan, Newe, Toal & Dooly, 2018). As both cameras took a picture of the same scene but at a different angle, an overlap of the scene points exists. In a next step, an algorithm engages in the process of stereo matching. This means that all the same scene points of the two images are identified. Subsequently, the horizontal displacement of the identified scene points is calculated. This horizontal displacement is called disparity. At last, the depth of the scene point is calculated from the known angle between the cameras and the disparity. If this process is done for all pixels in an image, a depth map, or in the case of Stereo imaging a disparity map, can be created (Ambrosch et al., 2010). The figure below offers a visualisation of the concept of Stereo imaging.

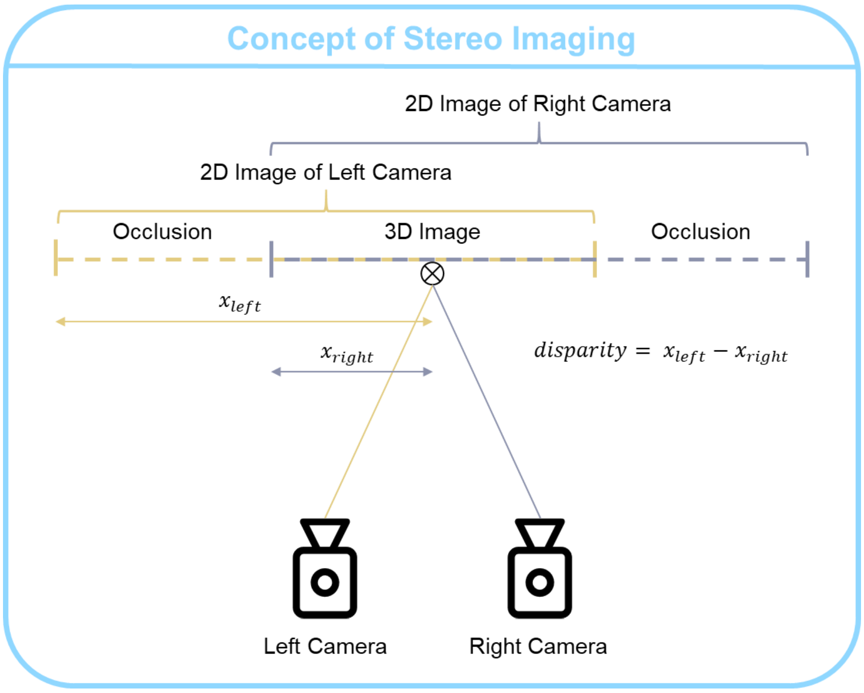


Figure 4. Concept of Stereo Imaging adapted from Kang and Liu (2014).

Two forms of Stereo imaging exist: passive and active. Passive Stereo imaging is the basic concept described above. Active Stereo imaging extends the passive concept by also including Structured Light imaging, as described in section 3.2 The advantage of active Stereo imaging is its higher depth accuracy and better performance when the camera or observed object is in motion (O' Riordan et al., 2018). The technology "active IR Stereo imaging" thus means that the concept of active Stereo imaging is applied with two infrared cameras. "Embedded Stereo imaging" refers to the passive version. In the sample of 3D/Depth cameras from section 2, this Stereo imaging version was typically implemented using two RGB cameras.

3.4 Comparison of Depth/3D-Sensor Technology

Each Depth/3D-Sensor technology has its strengths and weaknesses. There is no technology, which outperforms all others, instead each technology fits for a use case. This section elaborates some of the advantages and disadvantages, gathered from existing literature, with the aim to identify suitable depth/3D-sensor technology for a smart space application.

According to Gokturk, Yalcin, and Bamji (2004) a ToF system possesses several advantages over other types of depth imaging systems. A key advantage is the simple post processing calculation of the depth, whereas triangulation-based techniques like Stereo imaging require complex, resource intensive post processing algorithms. Also, from a cost perspective ToF is cheaper than other systems. Structured Light techniques require sufficient illumination contrast to guarantee visibility of the lighting pattern for the detector. Therefore, strong light sources are needed. ToF can operate with less strong light sources. ToF also possesses some drawbacks. The depth imaging system is affected by ambient light and secondary reflections, which can interfere with light rays sent out by the system (Boridy, 2019). The technology's z-axis resolution is rather granular, performing worse than Structured Light (Boridy, 2019). It was also observed that ToF has a worse performance when attempting to capture fast moving objects (Anderson, 2016). Overall, existing literature evaluated ToF as a robust depth sensing technology. It is best suited medium to long range measurements and to rapidly capture large objects. ToF can be applied both in outdoor and indoor use cases.

Structured Light depth imaging yields very precise results (Bell, Li & Zhang, 2016). Another advantage of this system is that it does not require scanning the surroundings, rather snapshots with the projected patterns capture the entire surrounding at once (Boridy, 2019). This provides a speed advantage, for example to analyse larger objects or larger spaces. This feeds into another benefit, the broad measurement range (Anderson, 2016). However, the resolution is very susceptible to ambient light. This drawback can be tackled by using a high intensity projector, at the expense of driving up the cost of the system. Another disadvantage are the resource intensive algorithms required to derive the depth map from the observed distortions (Boridy, 2019). Structured Light 3D sensing technology is best applied in use cases that require a high degree of accuracy on the z-plane. It offers a broad range and is thus suitable to quickly capture images of large objects. Due to its high susceptibility for ambient light, it best operates indoor in a controlled environment.

A benefit for both active and passive Stereo imaging is its cheap hardware setup. At minimum, only two regular 2D cameras are necessary (Boridy, 2019). The creation of the 3D image is then carried out by a computer algorithm. Those are available off-the-shelf and enjoy strong support, as the concept of Stereoscopy is mature (Boridy, 2019). Passive Stereo imaging has the advantage over other systems that it is not as negatively affected by ambient light. A key disadvantage of such a depth sensing system is its imprecision on the z-axis (Anderson, 2016), which active Stereo imaging attempts to improve. Moreover, the overall resolution of the created 3D image is low if simply two 2D images are combined. The final image does not have the same width, or field of view, as the two images do not have a complete overlap at the left and right boarders (Boridy, 2019). This is caused by the different angles from which the 2D images were taken. At last, the concept suffers from a low depth-of-focus (Anderson, 2016). The assessed literature found Stereo imaging to be an allrounder technology. The hardware setup for Stereo imaging is cheap. Furthermore, it experiences the least interferences with ambient light, compared to the other investigated technologies. This makes the depth sensing technology applicable in both indoor and outdoor settings. With its broad range it is suitable to cover larger objects, however, active Stereo imaging enables a decent 3D resolution in closer distance.

The analysed literature provided a general evaluation about the characteristics of the discussed technologies. The figure below attempts to provide an overview about three key characteristics of the analysed depth sensing technology: 3D resolution, operating distance/FOV Range, and indoor/outdoor usability. Those characteristics determine a first set of criteria to evaluate a use case fit in the subsequent sections.

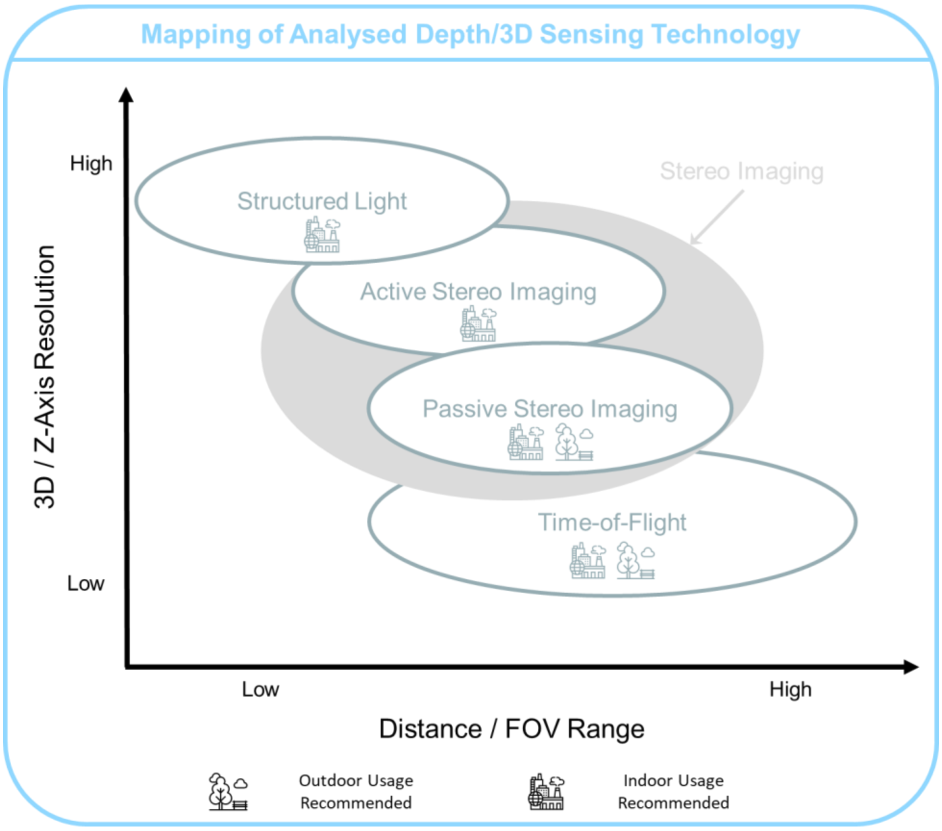


Figure 5. Mapping of analysed 3D/Depth-Sensing technology adapted from Anderson (2016).

However, the generalisations of the sensor technologies depicted in the figure above have to be treated carefully. They do not match in every aspect with the findings from the sample sensor device comparison. For example, the sensors Azure Kinect DK and Zed 2 delivered the second best and best 3D resolution. The sensor devices are employing ToF and embedded Stereo technology respectively. Literature generally evaluated those two types of technologies as the two with the lowest 3D resolution. This highlights the difficulty to evaluate one sensor technology as the best. There is no technology, which outperforms all others, instead each technology fits for a use case. Therefore, each device has to be evaluated based on its individual hardware configuration.

For a well-rounded analysis, a different object detection modality has to be considered. Due to the interest of the client in the depth-sensing capabilities of the Kinect DK, little thought has been given so far to explore the possibility to detect objects in 2D colour images with a normal RGB sensor. It is much cheaper and depending on the hardware, it can be more efficient in capturing frames (e.g. higher frame rate). According to Liu et al. (2020), most detection models, are based on 2D images. Much fewer models exist for identifying objects from its depth profile, as the minimum required depth resolution was only achieved in recent years (Verschae & Ruiz-del-Solar, 2015). So far, object detection in a depth modality finds most use in self-driving cars, autonomous aerial vehicles, and robotics (Murthy et al., 2020). All those areas of application employ depth sensing to detect obstacles and measure the distance to those for navigation purposes, not to identify the obstacles. Another key use of depth-sensing is to segment the image according to different depths, e.g. a foreground and background (Verschae & Ruiz-del-Solar, 2015). This adds new considerations for the use case fit. A possible application area are autonomous navigation systems. During image pre-processing, image depth segmentation could be used to cancel out background noise, which may distract the detection model. During image post-processing, the depth information can help to precisely extract the location of an object in an image. Such scenarios would require a depth sensor, otherwise an RGB camera would be a cheaper alternative.

4 Review of Related Academic Projects in the Field of Computer Vision

Academia and industry have a symbiotic working relationship. Academia serves as a pioneer to research and develops novel technology. Industry adapts the newly developed technology and seeks methods to commercialise it. Therefore, this analysis aims to identify key use cases in which academia has been successfully deploying 3D/Depth Sensing technology. Specifically, this review serves to provide an overview on what type of 3D/Depth Sensing technology academia has focused and in what kind of environments it was applied. This should enable to approach the main research question from another perspective.

4.1 Methodology of the Literature Review

A systematic literature review is conducted to minimise researcher bias and to ensure a high level of rigor in the different stages of the research. A systematic review requires a transparent and reproducible structure; thus, the methodology is built on the SALSA framework (Grant & Booth, 2009). It consists of a thorough search, identification, and appraisal of the search results, followed by a synthesis and analysis (Cipriani & Geddes, 2003). The databases Scopus, ProQuest, and IEEE were searched. To guarantee a broad and in-depth search, three search strings were construed out of the research questions. The search strings are assembled from the combination of the following sub-strings: computer vision, workspace, utilization, sensor technology, workspace, usage, and physical space. The exact search strings for each database can be viewed in appendix B.

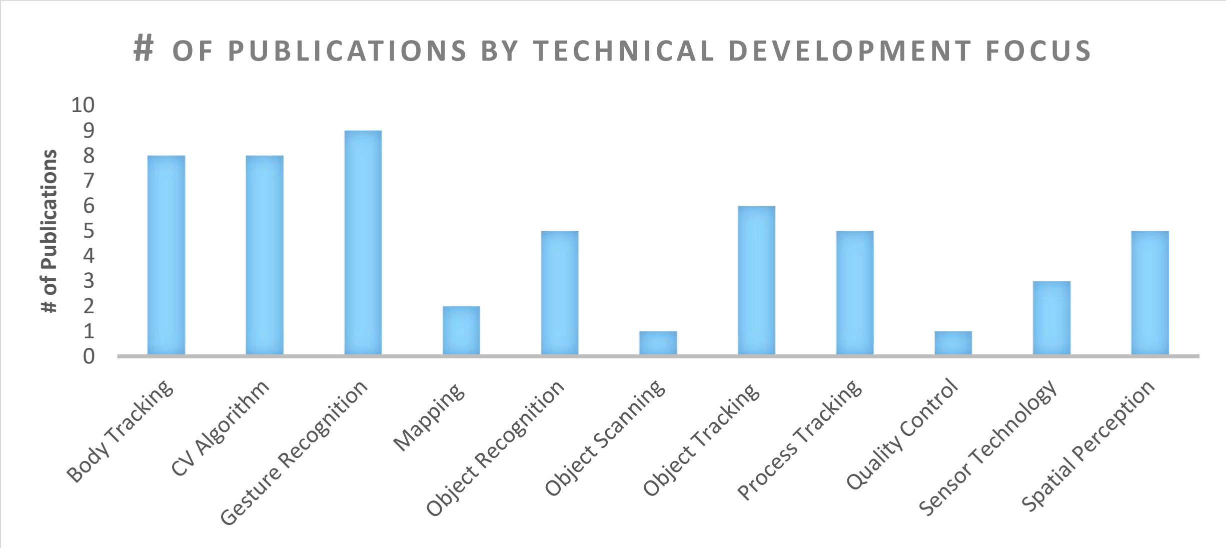
The search results were identified, appraised and synthesised in a funnel approach. First, the literature had to pass systematic inclusion and exclusion criteria. The criteria publication year, publication type, availability, and language posed the first eligibility stage. The rapidly evolving field of sensor technology and computer science demanded a recent publication time period; therefore, it was set between 2009 and 2020. After a search of just peer-reviewed academic journals did not yield a sufficient number of primary sources, working papers, dissertations, and scholarly magazines were also included. Allowing broader research criteria introduced the trade-off between resource quality and the anticipation of any literature gaps. Furthermore, the full text had to be available and written in English. As one of the aims is to identify new possible use cases of 3D/Depth sensing technology, the research scope was not further confined using specific criteria like research discipline or sector.

Subsequent to the initial identification of the sources followed a rigorous appraisal and selection. Integrated Excel sheets were used to catalogue the literature, which had been extracted in .csv files. First, a title followed by an abstract assessment was conducted. A source was assessed to be relevant for this investigation if it discussed at minimum a new technological development in the field of 3D/Depth sensing technology in combination with a use case. Of interest was also if the literature covered novel use cases applying existing depth sensing technology. Sources were not included if they solely discussed theoretical concepts of future developments in this field or advancements of 3D/Depth Sensing technology without any relevant use case. Appendix C exhibits an overview of the inclusion and exclusion criteria. By applying the mentioned criteria, reviewing the title and subsequently the abstracts, the sample set was reduced from the 1052 initially identified sources to 113. The full text review reduced the sample set further down to 54 sources. The grounded theory approach (Glaser & Strauss, 1967) was applied to synthesis and analyse the findings. First, those were thematically coded. Subsequently, they were compared and abstracted into groups of technical developments and use cases.

4.2 Discussion of the Findings of the Literature Review

An analysis of the metadata of the final sample literature provides key insights about academic's interest in the research field. Assessing the publication dates on a year-by-year basis, reveals an average rising trend. Between 2009 and 2012 only five sources of this sample were published. 2018 marks the peak with 12 publications (appendix D). The heightened interest in this research field is fuelled by the advancements in computer performance and its cheap and nearly ubiquitous availability. Cloud services and miniaturised local high-performance hardware enable the development of real-time applications. For example, one well-known use case is autonomous driving. An interesting observation offers an assessment of the publication type. With 45%, there are nearly as many dissertations in the sample as there are scholarly journals (appendix E). This underlines the novelty of this research area. In conclusion, the analysed temporal metadata points towards heightened research activity and a timely relevancy of this topic.

The synthesis and analysis of the sample literature, applying the grounded theory approach, yielded eleven abstracted groups with regards to the focus of technical development (appendix G). The assessment of those depth sensing technologies resulted in a grouping of 19 areas of application (appendix F). The graph below depicts the identified technological specialisations within the field of computer vision among the sample literature. In the subsequent paragraphs is each focus group highlighted in more depth if it satisfied the minimum threshold of five publications.



Graph 2. # of publication by technical development focus.

The literature sample indicated the greatest interest in gesture recognition. This research topic has been applied most frequently within Human-Computer- and Human-Robot-Interaction (HCI / HRI). HCI and particularly HRI receive an increasing attention, as it investigates means to increase the user experience and thus the adoption and acceptability of new computer related technology. For instance, Gang et al. (2018) applied gesture recognition to communicate with nursing robots. The researchers found that elderlies are much more receptive of the new technology if they can communicate with the robot using a familiar human-to-human communication style. Kwangtaek et al. (2016) stated that acceptance of gesture control can be even more enhanced if haptic feedback is incorporated in the communication loop. The concept of gesture-based control was also suggested to be abstracted and applied as a general-purpose multi-touchless interface (Krejov, Gilbert & Bowden, 2014). The current Covid-19 crisis gives this thought special attention, as such an interface could replace touch-based control panels in public spaces, e.g. for elevators or vending machines. Consequently, the transmission chain of germs, bacteria, or viruses via frequently touched surfaces is broken. So far, gesture recognition has focused predominantly on analysing the movements of hands, however, an important part of gestures are facial expressions. According to Zhang (2015), the analysis of facial expressions is challenged with regards to pose detection, face tracking, feature extraction, and model visualisation. Therefore, going beyond detecting reliably the six distinct human emotions (anger, disgust, fear, happiness, sadness, and surprise) would prove invaluable to HCI and create many novel use cases (Zhang, 2015). For example, detecting job stress could help prevent a negative economic impact on society of annually billions of US-dollars. Facial gesture detection applications provide a physically non-intrusive and non-invasive solution (Carneiro, Novais, Augusto & Payne, 2019). Combining such applications with ambient intelligence systems would create smart environments, which monitor stress levels of employees and actively adapt the surroundings to lower stress.

Body tracking applies similar concepts as gesture recognition but applies it to the entire body. It enables a cheap and passive collection of kinematic data, making it a highly interesting field of research for academia. Applications are found across a variety of use cases. This literature sample emphasised body tracking systems in the realm of medicine and worker wellbeing. Mbouzao (2013) employed a body tracking system as a cheaper, widely accessible, and passive sensing technology for a quantitative assessment of Rheumatoid Arthritis (RA). Kinematic data of a patient, potentially suffering from RA, is compared with a kinematic data of a healthy person for a simple but precise diagnosis. Moreover, a time series analysis of a patient’s kinematic data provides a mean to track the therapeutic effectiveness. Also, Scano et al. (2020) report strong potential of using body tracking technology for rehabilitation purposes. Their research tested RGBD cameras for those use cases and found this camera technology a promising alternative for impractical and expensive marker-based systems. Additionally, they envisioned body tracking applications as a means to monitoring task performance and ensuring worker safety. Seo (2016) researched precisely these use cases. The investigation consisted of building a body tracking application to measure the physical demands on construction workers. The purpose of measuring the physical demands is to increase workers’ safety and identify productivity and health issues. Computer vision technology enable novel methods to collect kinematic data, which outperform in efficiency and effectiveness. A similar study was conducted by Lun (2017) about healthcare workers. An alternative use case to analysing kinematic data for health purposes is to develop action recognition models out of it. This is a young research discipline and thus data set collection studies like Cai’s, Lu’s and Gao’s (2019) are still conducted. The researchers created a data set of tasks, which are performed by an employee at a work desk, video recorded from a first-person (egocentric) point of view. The intention is to promote the understanding of human activities of daily living (HADL), by offering a new perspective on hand-object interaction. This helps to conduct relevant cognitive and behavioural studies, next to efficiency analysis of workflows. Additionally, the data set lays the foundation to create new applications in the field of video surveillance and Human-Computer-Interaction.

The research paid an equal amount of attention to computer vision algorithms as to body tracking technology. This technology group did not explore specific use cases, instead it targeted to ease the development of new computer vision algorithms or improved existing concepts. For instance, Jacobsen (2014) developed the Smart Frame Grabber framework, which allows to cost effectively increase the performance of computer systems to run complex vision algorithms. Therefore, algorithms, which used to be non-real-time algorithms, can now generate outputs in real-time. Another focal point in this literature sample was the improvement of stereo vision algorithms. Dattagupta (2012) introduced a new method to reduce the errors around the edges of disparity maps in stereo matching algorithms. Contrastingly, Kaczmarek (2020) developed new stereo matching algorithms for stereo cameras using an Equal Baseline Camera Array (EBCA). As this system possesses a strong accuracy across a range of distances and performs in various light conditions, the system is an alternative to ToF, LIDAR, or Structured Light systems.

Object Tracking is the fourth most investigated research field of academia in this sample. Its application area is dominated by medical use cases for surgery. An optical tracking system is a key component of a surgery robot's vision system and enables several functionalities. For example, the tracking system supports the robot's navigation system (Oszust, 2018), or is the core part of the surgery tool localization system (Reiter, 2013). More advanced functionalities include fully automated assistive tasks, such as stitching a wound, which requires the tracking of surgical suture threads and surgery tools (Jackson et al., 2018). A major challenge in object tracking is the line-of-sight problem, due to occlusion (Wang, 2019). Various studies reported about this problem (Raza, 2013; Reiter, 2013; Oszust, 2018) and agree that a fusion approach delivers a satisfactory solution. For the surgery room this translates to a tracking system using multiple cameras mounted at different positions and a fusion algorithm to synchronise the data from the optical sensors (Wang, 2019). Raza (2013) proposed a different fusion object tracking framework, consisting of RFID and Computer Vision. The motivation for such a framework is that a system relying only on computer vision is computationally costly, ambiguous or even fallible, limited by occlusion, and in many cases not applicable in practice (Raza, 2013). Moreover, it improves accuracy and reliability of object recognition and tracking in 3D space. The study recognised a key use case for this fusion framework on a construction site to ensure workers' safety and enable activity analysis (Raza, 2013).

Academia's interest in object tracking is closely followed by object recognition. This is no surprise, as object recognition or at least detection is a prerequisite for object tracking. However, a surprise is that all research studies related to object recognition have been carried out between 2018 and 2019. This emphasises that there is a continued effort to improve and/or expand the application areas of object recognition. The advances in this research area will have follow-on effects for other research topics, such as object tracking or process tracking. Object recognition has a very broad application area. Exemplary use cases of this literature sample are yield mapping of fruit in agriculture (Roy, 2019), lane detection for autonomous driving (Xing et al., 2018), occupancy measurement in buildings (Choi, Quan & Cho, 2018), and robotic vision for indoor robotic applications (Sikdar, 2018).

Research into process tracking was equally prominent as research into object recognition and tracking. Process tracking technology experienced a boost with the availability of cheap depth cameras and advancements in object detection and recognition, as well as object and body tracking. Process tracking merges those technologies in order to model entire workflows. All of the analysed studies proposed use cases in industrial manufacturing, followed closely by construction projects. Furthermore, each paper recommended a fused system architecture, either by combining different sensor technology or by implementing a network of depth cameras. Edirisinghe (2019) reported that the first named process tracking architecture, consisting of technologies such as WLAN, UWA, GPS, Stereo or IR vision systems, are preferred on construction sites. The fusion models are designed for the purpose to monitor productivity and safety, as well as conduct ergonomic analysis of the construction workers. Workplace security is a reoccurring theme. For example, Rafibakhsh (2013) evaluated the Microsoft Xbox Kinect fit for a computer vision-based security system for large manufacturing jobsites. The other investigations placed a stronger focus on workflow tracking, modelling, and performance enhancement. The fusion model proposed by Bleser et al. (2015) stands out, as it collects visual data different to the other models, namely from an egocentric perspective. The system architecture consists of three body mounted RGBD cameras and two Inertial Measurement Units (IMU). However, also the works of Hogreve et al. (2016), as well as Faccio, Ferrari, Gamberi, and Pilati (2019) have particularities. The first named researcher group designed a multi-camera system out of Microsoft Xbox Kinects to not only track the assembly progress, but also to provide a gesture control system for the workers for their respective workstation. This significantly accelerated the work process, as workers could quickly enter necessary process data using gestures without having to go to the central computer terminal responsible for the workstation. The latter named group of researchers set out to replace traditional work measurement techniques for industrial production, which are inaccurate, slow, and require costly expert assessment. Over and above, the study determined relevant KPIs, drawn from the implemented passive observer system to analyse worker movements and task execution. The following KPIs were derived: duration to perform value-add and non-value-add activities, velocity profile and travelled distance of an individual worker and her/his hands, worker's walking path, occupied areas of the workstation, and component storage locations visited. The observer system itself consisted of a network of depth cameras, as well as a fusion algorithm to merge the different sensor streams and evaluate the quantitative KPIs.

Studies in the field of spatial perception was also of interest in this literature sample. Spatial perception refers to the ability of an agent to understand the environment in which it is situated (Mendez, 2017). Consequently, spatial perception is a daunting challenge in robotic vision and autonomous robots/(aerial) vehicles. Spatial perception entails three tasks, which enable a robot or some other vehicle to move autonomously through an unknown environment (Mendez, 2017). First the agent must create a 3D reconstruction of the environment. Subsequently, the agent has to plan a trajectory from its current location to the target location. This is also called path-planning. In the third step, the agent has to transfer its planned trajectory onto the 3D reconstructed environment and localise the path on real world maps. It has been suggested in the studies of Opoku (2013) and Shaoshan (2019) to support the localisation process by fusing information from additional sensors. For example, RFID tags, acting as markers, can be placed in indoor environments, as unique landmarks are often scarce (Opoku, 2013). The majority of analysed studies focused on improving one or more stages of the concept of spatial perception. Only

Roudaki (2013) attempted to apply spatial information to build more engaging and intelligent applications. Among others, the researcher prototyped an AR application, using 3D information about the spatial environment to generate better placements of AR projections.

4.3 Conclusion of the Literature Review

The systematic literature review sample detected 11 research areas in the field of computer vision over the past 10 years. Research topics, which received less attention in this sample, but remain indispensable, include developments in sensor technology, indoor and outdoor mapping, object scanning, and quality control. Partly, this can be explained as those technologies are further along the technology lifecycle trajectory, thus having lost the novelty factor. Nevertheless, this highlights the versatile use cases of computer vision applications. The figure below summaries the identified use cases in form of a word cloud. This figurative format visualises the frequency of the abstracted use cases of each technology in the font size.

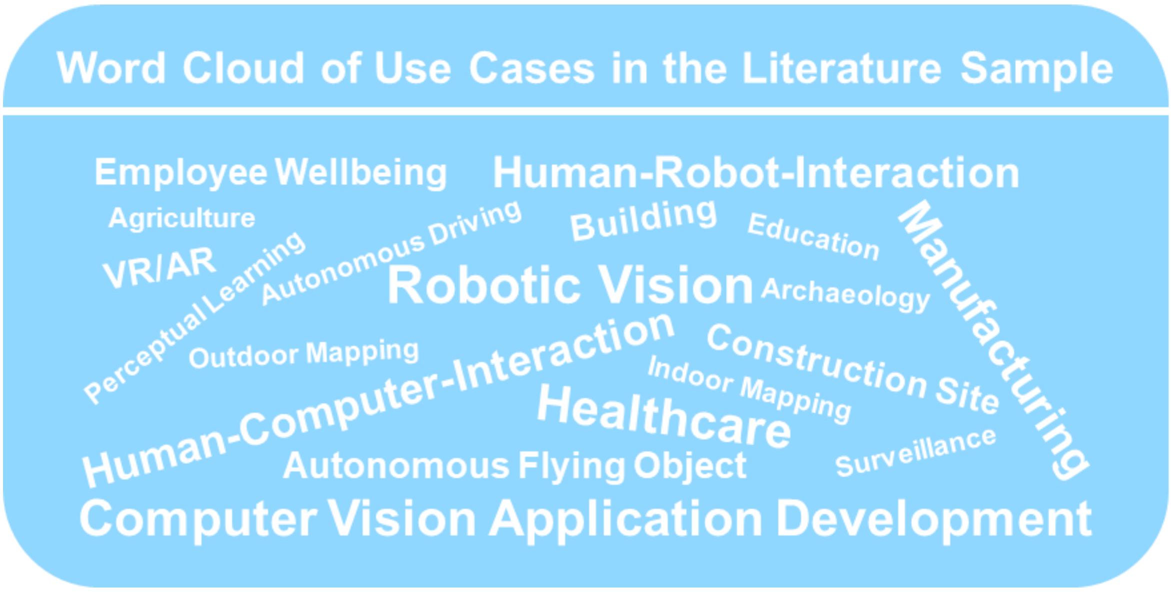


Figure 6. Word Cloud of use cases in the literature sample.

The findings of the literature review contribute to the investigation in three ways. First, it conveys a fundamental understanding about computer vision technologies. Secondly, the discussion of latest research indicates the current state of computer vision technology. This fosters the understanding of relevant and feasible use cases. At last, the identification of use cases set a starting point and guidance to theorise a use case fit for the Microsoft Azure Kinect DK sensor for different scenarios.

5 Business Case Analysis

The previous section provided a detailed answer for the first research question from an academic point of view, whereas the Business Case Analysis approaches the first research question from an industry perspective. This section attempts to identify attractive opportunities within the Smart Space/IoT market, which could be developed into use cases for the Azure Kinect DK. Additionally, this section focuses on providing a clear picture of the current market environments and key players of the Smart Space/IoT industry. This should tackle the third research question of this investigation.

5.1 Snapshot of the Smart Space/IoT Market

The Smart Space market was valued at $23.5 billion in 2018 and is forecasted to grow to around $86.5 billion by 2026 (Business Wire, 2019A). This resembles a CAGR of 17.6% between 2019 and 2026. The rapid growth in the Smart Space market has been driven by the strong upside potential, which smart spaces can offer. Wellner, Michalik, Manolian, and Griffin (2018) found that businesses, which employ smart sensor technology integrated into their facilities to engage in active building management, gain an edge over their competitors in key operational areas. Those include operational risk mitigation and operating cost savings, as well as employee attraction and retention. The Smart Space market can be divided by occupant or component. The latter named market division consists of the segments hardware, software, and services (Rake & Jagtap, 2019). Hardware makes up the largest segment, fuelled by the current need to upgrade buildings to smart spaces. The software segment follows closely, as it is required to orchestrate the sensors, analyse the data output, and derive actionable recommendations. The service segment is currently the smallest segment, but it is projected to become the most lucrative one (Rake & Jagtap, 2019). Once many buildings have been modernised to Smart Buildings, integrated sensors and the existing IT-infrastructure will need to be serviced. From an occupant perspective, the Smart Space market is split into residential, commercial, and industrial segments (Business Wire, 2019B). It is projected that the commercial segment is going to be the greatest contributor to the market growth over the forecasted horizon.

Many companies face increasing pressure from dynamic competitive rivalry across blurred industry boundaries, due to the ongoing development of smart, connected products (Porter & Heppelmann, 2014). Additional strain is exerted by the ongoing trend of sustainability. For example, many manufacturers and commercial companies have to meet certain environmental targets to satisfy their customer base (Parletta, 2019). Moreover, competition for top talent has been intensifying, requiring companies to transform their workplace experience into a sustainable competitive advantage to retain and attract talent (Roth, 2018). Smart spaces offer means to analyse the usage of one of the largest fixed cost factors of a company, office and factory space. It potentially allows a company to reduce resource wastage, optimise working processes, improve workers' safety and foster workers' productivity.

5.2 Analysis of Industry Solutions and Competitors

Despite the infant nature of the Smart Space market, hundreds of companies exist, which focus on offering IoT solutions across numerous industry verticals. The vast application areas of IoT technology combine previously separated industries. IoT requires hardware in form of sensors, controllers, gateways and communication modules. From a software perspective it utilises cloud computing, off-the-shelf AI offerings, and process modelling software. These components have to be combined with industry specific knowledge and client access to develop a well-formed IoT solution. Consequently, classical industry boundaries become blurry and the IoT players find themselves on a very competitive playing field. In order to achieve an in-depth analysis, the scope of the competitor analysis was limited. A geographic focus was laid on the two largest Smart Space markets, Europe and North America. Furthermore, industry reports from Markets and Markets (2019), Allied Market Research (Rake & Jagtap, 2019), Mordor Intelligence (2019), and Business Wire (2019A) were used as a guidance to identify key players in the market. This established a diverse sample of companies to colourfully depict the competitive landscape and current market solutions. The vetted competitors can be divided into multinational conglomerates, technology companies, and specialised service providers. Covering companies from different industry vertices and greatly varying maturities enables to observe different strategies to serve and capture the Smart Space market. Due to the ease of cross-selling existing IoT solutions to other industries, the entire IoT offering of the selected companies was compared. The figure below depicts the analysed companies and their organisational grouping.

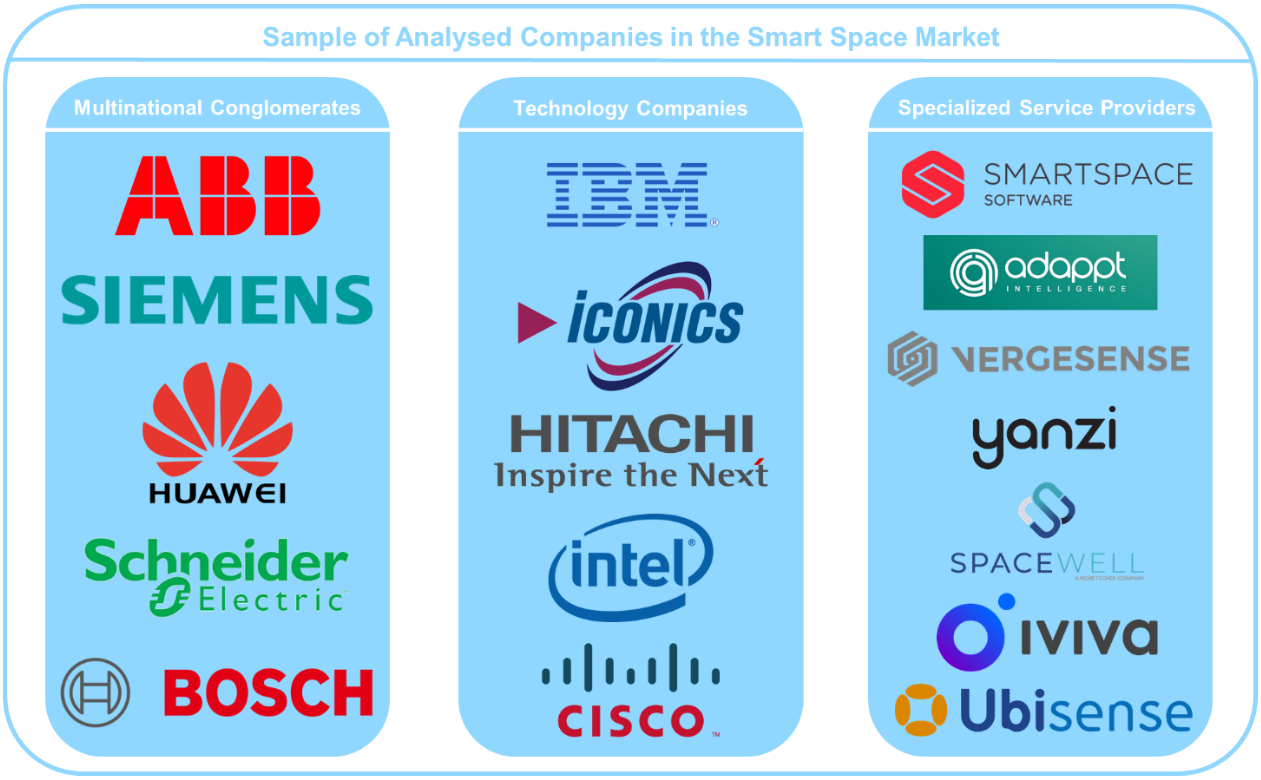


Figure 7. Organizational grouping of the analysed competitors.

The multinational conglomerates offer IoT products and services across a range of industry vertices. They particularly occupy the heavy industries such as Mining and Oil & Gas, as well as Manufacturing, Automotive, and Healthcare & Pharmaceuticals. A multi-layered strategy is observed among all of the players. The dominant goal is to develop a plug-and-play IoT platform. The momentary snapshot of the competitive field reveals a focus on Industrial IoT (IIoT) platforms. However, Bosch, Schneider Electric and Siemens also appear to offer sophisticated software to manage IoT devices for the commercial and residential real estate sector or mobility solutions. To ensure a wide reach of the platforms, some providers, like Bosch (2019), follow an open-source development approach, whilst others work with strategic partnerships. The second strategic aim is to retain as much of the IoT value chain in-house as possible. Despite the efforts to fulfil this aim, all conglomerates rely on third-party support due to lack of expertise or resources in one of the components to offer fully integrated solutions. ABB, Schneider Electric, Siemens, and Bosch either partnered with a cloud and/or technology company to supplement their cloud offering or to support their analytics service with off-the-shelf AI solutions. Currently, the major use cases for the multinationals are in the field of asset management, predictive maintenance, and manufacturing process optimisation. Particularly, the latter named service receives considerable attention by advancing the technology to more precisely map the manufacturing floor/process to a digital twin. The versatile applicability of IoT solutions across the diverse fields of discrete and process manufacturing feeds into the conglomerates' core advantage; a large client pool in which nearly every customer is a target. The ease of applying almost the same IoT technology with only minor modifications to other domains creates a strong cross-selling potential in a conglomerate structure. Consequently, the conglomerates have an advantage in the speed to market in industries, which previously did not enjoy an IoT offering. Looking at each of the companies' more closely, Huawei stands out for two reasons. Unlike its competitors in the sample segment, Huawei can offer all IT infrastructure, cloud and AI computing services itself. However, it is the only company that relies entirely on third-party sensor technology, other than communication modules (Huawei, 2019). Secondly, the company has extensive knowledge in computer vision (Huawei, 2020). This form of data collection is not as prominently offered by the others. Schneider Electric's EcoStruxure IoT platform scores with its specialized use cases across five industry domains (Schneider, 2020). This places Schneider as a leading company for off-the-shelf advanced IoT applications interoperable in its IoT ecosystem. Bosch rolled out its *Bosch IoT Suite* later than its competitors in the sample, nevertheless, it offers a very competitive solution (Bosch, 2019). It strongly benefits from Bosch's sensor technology and is similar to Schneider Electric's platform. Noteworthy are Bosch's IoT solutions in the food production value chain, in which it successfully combines machine vision, GPS, and other sensor data for the optimal cultivation (Bosch, 2020). Siemens' MindSphere has strong potential to become a leading IoT platform. MindSphere has not only incorporated the benefits of its competitors but is additionally supported by Advanta (Siemens, 2019). Advanta is a specialized division, which focuses entirely on consulting and implementing IoT projects across eight different industry verticals. Thereby, it applies the full range of available sensor technology.

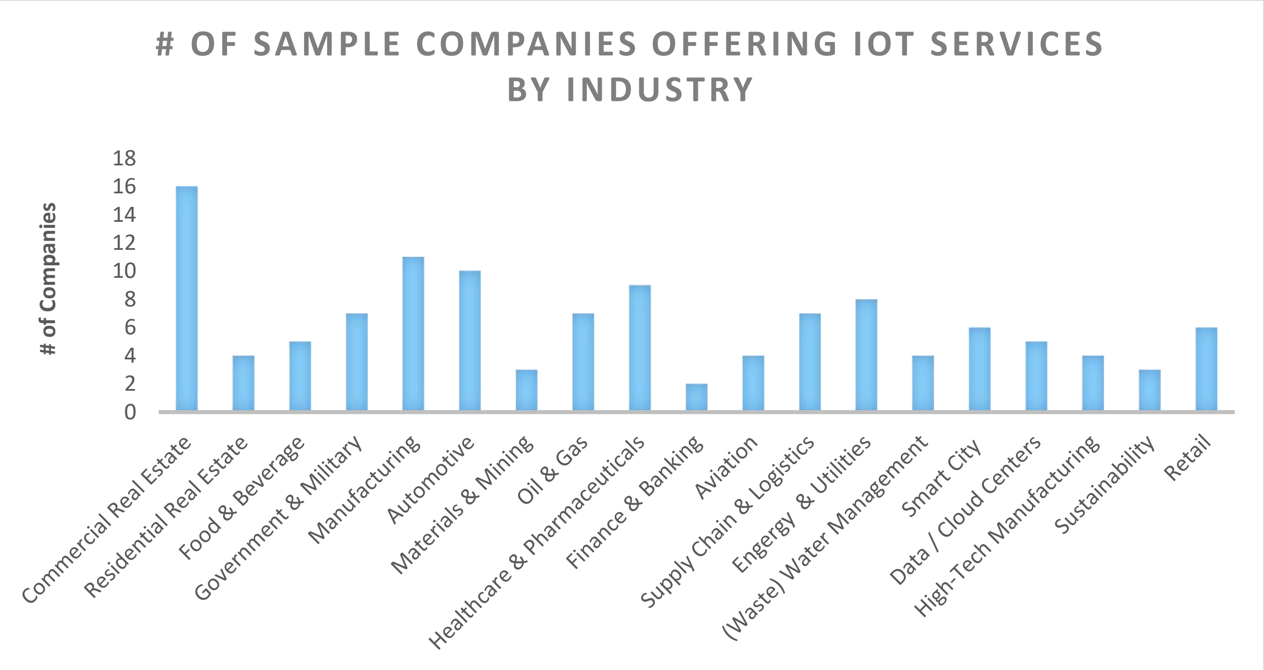
It consists of 500 consultants and 7500 developers and engineers, whose goal is to establish MindSphere as the status quo IIoT platform (Siemens Advanta, 2020). ABB has a smaller but similar offering to Siemens. It serves fewer industry verticals and is even more focused on heavy industry (ABB, 2019 & 2020).

The technology company sample segment consists of players, which have focused on SaaS offerings for the IoT space. The degree of industry specialization varies greatly. Three overarching strategies could be observed among the technology companies. One strategy foresees the development of industry vertical specific SaaS offerings. The other strategy pursues the goal of investing resources into the development of domain-specific solutions for the IoT. These solutions are not industry specific but rather offer a general tool/developer kits for other companies to use or to be applied in a joint consulting and implementation project with the technology company. Exemplary solutions are data management and processing platforms in the form of cloud services or data analytics suites. Those enable to tap into machine learning software such as computer vision, and cyber security applications for IoT communication and data processing. While the first strategy aims to establish an enclosed SaaS ecosystem, the other strategy attempts to establish a platform of SaaS IoT services applicable to different stages of the software value chain. The third strategy is a mix of the first and second strategies. Iconics, a SaaS for IoT software, embraces the first strategy. The company currently offers 12 different automation software solutions across 12 different industry verticals. Hitachi, Cisco, and Intel are the only technology companies in this sample that manufacture some hardware or sensor equipment for IoT devices. Those three companies are settled somewhere in between the two discussed strategies. Similar to Iconics, Hitachi strongly orientates itself along selected industry verticals with its IoT analytics platform Lumada. However, its platform serves fewer industries than Iconics and is more focused on analytics software with a smaller offering for system management (Hitachi, 2019). An area in which the company leans more towards the second strategy is its Video Intelligence service. This service enables customers to utilize non industry specific computer vision software to develop customized applications (Hitachi, 2020). Intel and Cisco are more positioned to follow the second strategy. Both companies offer SaaS IoT applications for device management and analytics. From a hardware perspective, they manufacture gateways and other connectors to facilitate communication within an IoT network (Cisco, 2019; Intel, 2019). Intel, as a chip manufacturer, also specialized on producing chips specifically for the use of IoT devices. The chip maker further stands out for its progressive SaaS platform Open Vino to ease the development of deep-learning machine vision applications (Intel, 2019B). The SaaS offering is supported by its own depth-camera series RealSense. With the creation of the subsidiary Mobileye, Intel infiltrates the Mobility industry by leveraging its expertise in machine vision and exploring new fields of computer vision. Mobileye may serve to further advance Intel in the area of smart vision applications for IoT (Intel, 2019B). Contrastingly, Cisco’s Meraki offering consists of more basic smart vision applications, such as security, object detection, or motion heatmaps (Cisco Meraki, 2020). IBM fully embraces the second strategy. With its Cognitive and Cloud services, it helps companies to realize their desired IoT solutions (IBM, 2019). Similar to Intel, IBM has computer vision tool kits within its Watson product offering, but those were not tailored to IoT use cases. However, IBM is also extending their offer in full end-to-end SaaS solutions for industry verticals with its platform IoT Watson (IBM, 2019). Therefore, there are early signs that a domain specific IoT platform, combined with a set of industry specific applications, emerges to become the dominant strategy.

The specialised service provider sample segment is characterised by young companies, which only operate within a single industry vertical. Six out of the seven service providers focus on commercial real estate. This industry sector can be split into the sub-segments hospitals, airports, hospitality, shopping centres, and office space. This entire sample segment pursues a niche strategy. Within its industry vertical, it specialised on the provision of a specific service. For instance, unlike Iconics' SaaS offering across various industry verticals, Ubisense specialises on discrete, job shop, and batch manufacturing. Its core focus is to provide real-time location systems. By creating a digital twin of the manufacturing process and using ultra-wide-band sensors and tags, the system traces movement of tools, workers, and production input (Ubisense, 2020). Ubisense uses this data to then optimise manufacturing processes. Office space dominates the commercial industry vertical. Yanzi, Adappt, and Vergesense work on smart office space. All three companies manufacture their own smart sensors. Yanzi and Adappt provide sensors, which are able to measure temperature, humidity, and motion. Furthermore, Yanzi developed sensors to monitor air quality and ambient noise (Yanzi, 2020). Adappt (2020) and Vergesense (2020) both developed a visual sensor for occupancy measurement with low resolution pixels. All three companies offer space management and analytics software. Yanzi differentiates itself from Adappt and Vergesense in terms of its strategy to allow compatibility with third party systems and services. Adappt (2020) and Vergesense (2020) only allow third-party integration on two levels. Third parties can use Adappt's and Vergesense's sensors for their IoT platform or third parties are able to integrate some of Adappt's and Vergesense's IoT management and analytics services into their software through open APIs. Contrastingly, Yanzi (2020) offers a completely open end-to-end IoT solution. Its sensor can be used in third-party IoT software, Yanzi's IoT system architecture provides open APIs to run third party IoT platforms or services, and Yanzi's IoT platform accepts third-party sensors. In contrast to the previously discussed companies, SmartSpace, SpaceWell, and Iviva (by Eutech) are pure SaaS providers. They accommodate customers across several of the commercial real estate sub-segments. Both companies are providing facility management, as well as space and process analytics software (Manifold, 2020; SmartSpace Software PLC, 2019; Eutech, 2020). However, they are reliant on compatible third-party sensors.

5.3 Findings of the Business Case Analysis

Assuming a bird's eye perspective on the entire sample allows to pull together all observations made in the segments. Nearly all companies offer some kind of solution for the Commercial Real Estate (CRE) space. This attention is in part driven by CRE's market size of $3.2 trillion in revenue in 2018, making it the fifth largest industry in the world (IBISWorld, 2020). Moreover, the fact that CRE had not experienced any greater technological advancements, made the industry a blue ocean for IoT technology. This might explain the majority of startups in the sample segment focusing on CRE. Other industries, which enjoy a strong IoT offering are the Manufacturing, Automotive, Healthcare & Pharmaceuticals, and Energy & Utilities industry. Just like the CRE industry, all those verticals have an extraordinarily large market size in common. The currently weak Retail industry might become an IoT technology *forerunner*. Due to its industry nature, it has to rely on passive IoT systems, which observe the shopper. Such IoT technology may offer wider applicability in other industries to optimize working processes, which cannot be tracked with simple tags. Instead smart vision technology could passively observe the processes at a lower cost and fewer physical restrictions. The graph below depicts the sample companies' IoT offering across the various industries.



Graph 3. # of sample companies offering IoT services by industry.

Mapping the sample companies according to their sensor and software offering yields key insights. None of the companies offered just IoT management software. Instead all firms provided at least analytics services. The next step in terms of software offering is to use the data analytics to generate actionable recommendations for process optimizations. Only a few companies are actively engaging in this task. It must be noted that not every company that offers process optimization solutions delivers them through their IoT software. Some companies, like IBM, deliver those solutions via a combination of their advanced analytics software and consulting services. A clear pattern can be observed that companies specialize in IoT solutions for distinct use cases within an industry vertical. Those systems have the capability of including process optimization applications. Abstracting this observation, there is a general trend emerging to upgrade pure analytics software into process optimization applications. Analyzing the sensor offering of the sample, two distinct groups become apparent. Some companies operate only as a SaaS provider depending on compatible third-party sensors. The other group of companies pursues the strategy of manufacturing their own IoT sensor devices, which are compatible with third-party IoT platforms, as well as providing an IoT SaaS. Only one firm is in danger of being stuck in the middle, Hitachi. The competitor analysis identified Schneider Electric, closely followed by Siemens, as an outpacer. The company built an IoT platform with specialized software ecosystems for five industry verticals. The SaaS offering is supported by a broad sensor offering. No other company in the sample could match Schneider's IoT solutions in terms of its breadth and depth. The figure below displays the mapping of the sample companies according to their IoT capabilities.

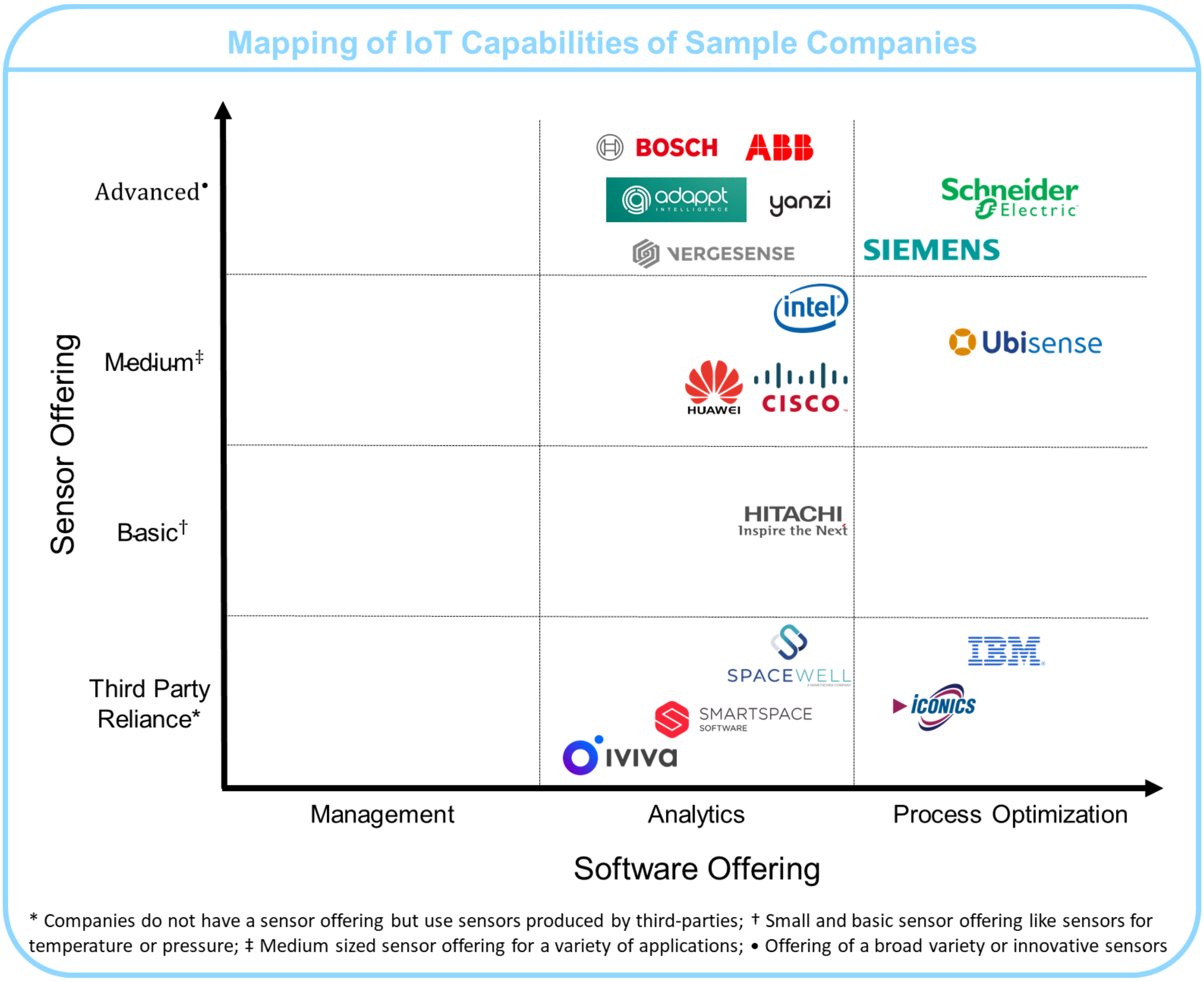


Figure 8. Mapping of IoT capabilities of sample companies.

6 Recommendations of Business Use Cases for the Microsoft Kinect DK

Central questions of the background research and the investigation were to evaluate the Kinect DK against competitor sensors and to suggest suitable use cases for this new depth-camera. Section 2 addressed the first named question. The latter question was tackled by section 3, 4, and 5 from a technical, academic and business perspective respectively. These points of views are now attempted to be combined in order to provide a final set of recommendations and the input to tackle the second deliverable of this project. The Microsoft Azure Kinect DK has several strengths, but also weaknesses, as observed in the technical evaluation. Consequently, the right use case fit must be determined to play to the strengths of the Kinect.

The Kinect DK is a multi-sensor passive observer. The Kinect's 12-MP RGB camera enables capturing of high-resolution images for object and human detection and identification. The 1-MP depth sensor, using ToF technology, offers a decent resolution at a medium operational range (Microsoft, 2020). Moreover, this depth sensing technology can be applied in- and outdoors. The depth-sensor in combination with the RGB camera empowers the Kinect to provide accurate data for image depth segmentation, motion tracking and identification. This passive observer technology has a key advantage in the IoT realm. Over time it is less expensive than maintaining other sensor technology, such as tags and a detector. Additionally, in some processes there may be physical constraints to use tags. However, a central drawback is the issue of occlusion. The Kinect has a workaround, by having integrated sync pins. Those pins allow to conveniently build a Kinect camera network, by automatically synchronising their data streams (Microsoft, 2020). The IoT sensor is further equipped with a gyroscope, accelerometer, and a 7-microphone array. The first two named sensors generate data for spatial tracking and sensor orientation (Microsoft, 2020), making the Kinect attractive for robotic vision and navigation. The microphone-array provides the ability for wide-range sound capture (Microsoft, 2020), valuable for a voice controller. At last, the Kinect DK profits from its interoperability with the Azure ecosystem. Its data outputs are aligned with the input requirements for the Azure Cognitive Services (Microsoft, 2020). This makes the Kinect very attractive from a development perspective, as services like AI object and face detection models or speech recognition models can be directly accessed via the cloud. The Kinect DK is a sensor fusion model in itself, therefore use cases that require the sensor's versatility are of greatest interest. Five areas of application have been identified, which offer such use cases. Those are proposed below.

6.1 Office Space

Up until recently, office space has been regarded by companies as an essential but costly and passive asset. Other than providing a space, it was not perceived to add any additional value. Despite being the second or third largest fixed cost position for many companies, office space utilisation was rarely actively being tracked. New IoT technology can help office space to become an active asset, contributing to the employee wellbeing, productivity, and effective space utilisation. The competitor analysis revealed that particularly young companies attempt to solve problems around ambient management and occupancy measurement. The latter problem has not been solved optimally yet. Current solutions often use light sensors attached under the table to detect desk occupation in an unassigned seating environment. This often results in false readings, as people go to meetings or into breaks. Instead, the Kinect could be employed to capture images of a desk and derive an occupation probability, based on the detected objects left on the desk. Moreover, entire floors and rooms can be observed to capture foot-traffic data. Room usage heat-maps can be derived from this data input, providing valuable information for the post Covid-19 working environment to detect potential spreading hotspots in need of extra ventilation or disinfection. The Kinect's capabilities can be further exhausted by engaging in workstation and worker wellbeing analytics. The tasks performed by an employee can be tracked and analysed, to allow for an optimised seating plan. Stress levels, as well as posture and distance-to-screen can be monitored using the Kinect's depth-sensing and RGB camera for motion and facial analysis. To improve the accuracy of the detection model's predictions, any background disturbances could be minimized through depth segmentation. Such applications have the potential to improve the employees' health through real-time feedback (Carneiro et al., 2019).

6.2 Industrial Manufacturing and Construction Sites

Since the existence of industrial manufacturing, companies have been engaging in continuous process optimisation. However, human executed tasks have always posed a challenge to gauge. Those require non-intrusive and non-invasive measurement techniques to not obstruct the workflow. Faccio et al. (2019) tested positively a multi-camera system to capture motion data. It enabled the derivation of much more accurate and insightful KPIs, compared to the traditional ones. The Kinect DK possess the exact combination of functions to commercially develop such a system. The competitor analysis showed that only a handful of companies are currently offering such a system, allowing a smoother entry into the highly contested but large IIoT market. Making an assembly station interactive and responsive would be the next level (Hogreve et al., 2016). A worker's motion data combined with object tracking and detection could be used to determine the assembly progress. A workstation could actively respond to this information. Further, the Kinect serves as the ideal workstation controller. Gesture-recognition and voice control could be employed by the worker to control the active elements of an assembly station. This would result in higher productivity and a workload relief for the worker.

The Kinect DK's technology could also be employed on construction sites, as the ToF depth-sensing technology also performs well outdoors. Similar as in industrial manufacturing, a camera-network could be installed for large scale construction projects for automated progress tracking and process optimisation. Academia also proposed the development of a work safety system based on computer vision. It could detect and register workers, which are located in hazardous areas of the construction site (Raza, 2013; Edirisinghe, 2019). The business analysis revealed little competition in this field, making it attractive for a niche entry.

6.3 Retail

The Kinect's ability to generate useful image data for object and motion detection, tracking and identification at a medium range distance proves valuable for retail floor analysis. A foot-traffic analysis could reveal which aisles are most frequented and what the preferred routes through a shop are. Furthermore, it could track shelf-inventory to prevent costly stockouts. Motion identification could be used to identify in front of which products shoppers spend the most time. Detecting thinking poses or the motion of picking or placing back a product provide invaluable insights. The latter named motions would also be central to develop cashier-less stores. Those data points would enable brick and mortar stores to develop an enhanced shopper experience. They would also support decision making for product placement and identifying prime shelf space by target groups to determine dynamic shelf space rental prices.

6.4 Medicine

The Kinect's integrated gyroscope and accelerometer enable sensor orientation. This is essential in combination with the RGB and depth sensor if the Kinect is mounted onto a robotic arm as the robotic vision unit. Additionally, the 7-array microphone would enable reliable voice control. Having all sensors in one device, eases the development as data streams are compatible and bundled. Such a vision unit, which can simultaneously act as a voice controller, could find particular use cases for (assistive) surgery robots (Reiter, 2013; Jackson et al., 2018). Surgeons could gain a steady helping hand, which is conveniently controlled via the surgeon's voice. Also, all sensors of the Kinect would be utilised in this multi-requirement use case. However, depending on the surgical task, an investigation whether the Kinect's ToF depth-sensor has sufficient accuracy, would be necessary. The Kinect's motion tracking ability would also prove valuable for (at home) rehabilitation training, as proposed by Scano et al. (2020). The sensor's data output is rich enough for AI models to check for posture and exercise control.

7 Legal and Ethical Implications of the Proposed Use Cases

The greatest threat to any of the proposed use cases is not necessarily a bottleneck in the technology, but rather legal and ethical reasons. The use cases for the office space, for industrial manufacturing, and for the construction site require observing the workstation/place of work, including the employee/worker, at a high-fidelity resolution. Consequently, this program would be greatly intruding the privacy of the employee, as identification is possible. From a legal perspective, such a program is prohibited in certain countries, such as Germany (Böhm & Ströbel, 2017), where employers are not allowed to surveil their employees. Moreover, the project raises several ethical questions. First and foremost, to what extent is it justifiable that an employer is violating the privacy of the employees for the benefit of higher productivity? There are obviously many arguments which would support the request, because the higher productivity will help the overall financial performance of the company and makes it more competitive. It will strengthen the company to introduce better tools, rules and organizational structure to support the wellbeing of the employees and potentially allow the company to grow and create new jobs. The growth in economic value reflects in gains of the social value. On the other hand, if no growth opportunities exist, it may pave the ground for job reductions as less employees will be able to deliver the same amount of work. Concerning could be that it will create transparency around individuals' behaviors and habits, which may allow conclusions on the health situation of the individuals (how often do they need to use the bathroom, are they very tired, is the employee frequently pausing etc.). The intrusion of an employer into the employee's privacy may even create the opposite of the envisioned increase in productivity and wellbeing. The project might also fail because employees start performing worse. The constant feeling of being observed and monitored, could result in employees feeling pressured (Anteby & Chen, 2018). This in turn may have negative effects on the mental health, which therefore reduces the productivity of the employees (MacRae & Murphy, 2017). Also, many employees might refuse to work in an environment, in which they are constantly observed.

8 Conclusion and Limits of the Investigation

8.1 Conclusion

The investigation provided an in-depth answer for the three central research questions. Industry analysis revealed that the IoT industry is rapidly growing with strong competitive rivalry among its key players. The dominant goal among the sample companies is to develop a plug-and-play IoT platform. Some companies pursue a purely software-centric strategy to achieve this goal. Others aim to maintain the sensor development in-house along with software solutions. Several trends were also identified. Most notably, the majority of the sample companies focus on the IIoT. Within the space of IIoT, the players attempt to move from analytics services to offer systems for process optimisation. Consequently, many newly developed IIoT solutions specialize on an industry vertical. The findings centred around the Azure Kinect DK showed that the IoT device is a highly competitive depth sensor. Within the competitor sensor sample, it is positioned as a quality leader, scoring the second highest in terms of functionality. Furthermore, the competitive product mapping highlighted that the Kinect is a device, which excels in use cases requiring several sensor functionalities. The core research question regarding the identification of potential use cases for the Kinect DK was answered with five recommended use cases. These have been developed over the following areas of applications: office space, industrial manufacturing and construction sites, retail, and medicine. The recommended use cases are the key input for the second deliverable of the project. In collaboration with Avanade, one of the use cases is selected and a proof of concept application implemented to assess its feasibility as a potential new service for Avanade's clients.

8.2 Limitations

The findings of the investigation have to be viewed in light of the investigation's limits. Due to the novelty of the research topic, the paper had to consider a variety of different types of non-peer reviewed and non-academic sources. Additionally, the sources had to be accessible. Consequently, source types range from industry and company reports, to journal articles, PhD dissertations, and blog posts. This impairs the resource quality but ensures a sufficient coverage. To minimize this effect, each source was thoroughly inspected and subject to inclusion and exclusion criteria, where applicable. In absence of time constraints, the research quality could have been improved by interviewing industry experts. This would have provided insider knowledge about the current industry sentiment and the applicability of depth-sensors for commercial applications. Potential researcher, reporting, and confirmation bias pose a further limit to the findings of the investigation. Certain parts of the analysis, such as the systematic literature review or the competitive product mapping, required non-guided judgements. These were attempted to be minimized and where possible backed by judgement criteria or quantitative data. Moreover, the researcher always aimed to maintain high levels of integrity throughout the study. Also, it must be considered that similar biases could potentially be passed on from the included sources themselves. Another limitation of the findings is that the suggested use cases have not been validated by the respective target customer groups. Surveys could have been sent to the target groups to understand their needs and pain points in greater detail. This would have helped to identify, which recommended use case has the greatest success potential to become a product, which the target customers actually want.

Bibliography

Note, the following bibliography is cited according UCL (IOE)'s Harvard citation style.

ABB (2019). *Annual Report 2019* [Online]. Available at: https://search.abb.com/library/Download.aspx?DocumentID=9AKK107680A2993&LanguageCode=en&DocumentPartId=&Action=Launch [Accessed: 21 June 2020].

ABB (2020). *Industrial IoT Applications* [Online]. Available at: https://new.abb.com/control-systems/features/industrial-IoT-services-people-use-cases [Accessed: 21 June 2020].

Adappt (2020). *Home* [Online]. Available at: https://adappt.com/ [Accessed: 23 June 2020].

Ambrosch, K., Humenberger, M., Olufs, S. & Schraml, S. (2010). 'Embedded Stereo Vision' [Online]. *Smart Cameras*, Boston, MA: Springer US. pp. 137–157, doi:10.1007/978-1-4419-0953-4_8 [Accessed: 2 July 2020].

Anderson, J. (2016). 'A Closer Look at 3D Imaging' [Online]. *Quality Magazine*. Available at: https://www.qualitymag.com/articles/93544-a-closer-look-at-3d-imaging?v=preview [Accessed: 3 July 2020].

Anteby, M. & Chan, C. K. (2018). Why Monitoring Your Employees' Behavior Can Backfire. *Harvard Business Review* [Online]. Available at: https://hbr.org/2018/04/why-monitoring-your-employees-behavior-can-backfire [Accessed: 25 June 2020].

Asus (n.a.). *Xtion PRO LIVE | 3D Sensor* [Online]. Available at: https://www.asus.com/3D-Sensor/Xtion_PRO_LIVE/specifications/ [Accessed: 25 June 2020].

Bell, T., Li, B. & Zhang, S. (2016). 'Structured Light Techniques and Applications'. *Wiley Encyclopedia of Electrical and Electronics Engineering* [Online], pp. 1–24, doi:10.1002/047134608X.W8298 [Accessed: 2 July 2020].

Böhm, W. T. & Ströbel, L. (2017). New Case Law on Restrictions for Employee Monitoring in the Workplace in Germany. *Employment Privacy, International/EU Privacy* [Online]. Available from: https://www.hldataprotection.com/2017/08/articles/international-eu-privacy/new-case-law-on-restrictions-for-employee-monitoring-in-the-workplace-in-germany/ [Accessed: 25 June 2020].

Boridy, R. (2019). 'Pros and Cons of Popular 3D Technologies' [Online]. *Novus Light*. 25 February. Available from: https://www.novuslight.com/pros-and-cons-of-popular-3d-technologies_N8910.html [Accessed: 3 July 2020].

Bosch (2019). *Annual Report 2019* [Online]. Available at: https://assets.bosch.com/media/global/bosch_group/our_figures/pdf/bosch-annual-report-2019.pdf [Accessed: 21 June 2020].

Bosch (2020). *Internet of Things* [Online]. Available from: https://www.bosch.com/internet-of-things/ [Accessed: 21 June 2020].

Business Wire (2019A). '$86+ Billion Smart Space Market 2019-2026: Analysis by Component, Space Type, Application, End-user and Region - ResearchAndMarkets.com.' [Online]. Available at: https://www.businesswire.com/news/home/20190925005419/en/86-Billion-Smart-Space-Market-2019-2026-Analysis [Accessed: 15 June 2020].

Business Wire (2019B). 'Global Smart Space Market Report 2019: Technologies and Applications Analysis & Forecasts 2018-2024' [Online]. Available at:

https://www.businesswire.com/news/home/20191217005383/en/Global-Smart-Space-Market-Report-2019-Technologies [Accessed: 9 June 2020].

Center for Applied Statistics Courses UCL (2018). *Ranking Data* [Online]. Available at: https://www.ucl.ac.uk/child-health/short-courses-events/about-statistical-courses/research-methods-and-statistics/chapter-8-content-0 [Accessed: 27 June 2020].

Cisco (2019). *2019 Annual Report* [Online]. Available at: https://www.cisco.com/c/dam/en_us/about/annual-report/cisco-annual-report-2019.pdf [Accessed: 21 June 2020].

Cisco Meraki (2020). *Cloud Managed Smart Cameras* [Online]. Available at: https://meraki.cisco.com/products/smart-cameras/ [Accessed: 23 June 2020].

D'Aveni, R. A. (2007). Mapping Your Competitive Position. *Harvard Business Review* [Online]. Available at: https://hbr.org/2007/11/mapping-your-competitive-position [Accessed: 27 June 2020].

e-con Systems (2016). *Tara USB 3.0 Stereo Camera* [Online]. Available at: https://www.e-consystems.com/3D-USB-stereo-camera.asp#key-features [Accessed: 25 June 2020].

Eutech (2020). *Smart Environments that make you smile* [Online]. Available at: https://www.iviva.com/smartspaces/ [Accessed: 23 June 2020].

Gokturk, S. B., Yalcin, H. & Bamji, C. (2004). 'A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions'. *2004 Conference on Computer Vision and Pattern Recognition Workshop*, 2004(January), p.35. [Accessed: 21 June 2020].

Hitachi (2019). *Hitachi Integrated Report 2019* [Online]. Available at: http://www.hitachi.com/IR-e/library/integrated/2019/ar2019e.pdf [Accessed: 21 June 2020].

Hitachi (2020). *Video Intelligence* [Online]. Available at: https://www.hitachivantara.com/en-us/products/video-intelligence.html [Accessed: 23 June 2020].

Huawei (2019). *2019 Annual Report* [Online]. Available at: https://www-file.huawei.com/-/media/corporate/pdf/annual-report/annual_report_2019_en.pdf?la=en [Accessed: 21 June 2020].

Huawei (2020). *Huawei-HDC-Computer-Vision-Plan* [Online]. Available at: https://www.huawei.com/en/news/2020/3/huawei-hdc-computer-vision-plan [Accessed: 21 June 2020].

IBISWorld (2020). 'Global Biggest Industries by Revenue 2020*'* [Online]. Available at: https://www.ibisworld.com/default.aspx [Accessed: 24 June 2020].

IBM (2019). *2019 Annual Report* [Online]. Available at: https://www.ibm.com/annualreport/assets/downloads/IBM_Annual_Report_2019.pdf [Accessed: 21 June 2020].

Iconics (2020). *Internet of Things Solutions* [Online]. Available at: https://iconics.com/Solutions/IoT [Accessed: 23 June 2020].

Intel (2019B). *2019 Annual Report* [Online]. Available at: https://s21.q4cdn.com/600692695/files/doc_financials/2019/2019-Annual-Report.pdf [Accessed: 21 June 2020].

Intel (January 2019A). *Intel RealSense D400 Series Product Family Datasheet* [Online]. Available at: https://www.intel.com/content/dam/support/us/en/documents/emerging-technologies/intel-realsense-technology/Intel-RealSense-D400-Series-Datasheet.pdf [Accessed: June 25, 2020].

Intel (June 2020). *Intel RealSense LiDAR Camera L515 Datasheet* [Online]. Available at: https://www.intelrealsense.com/lidar-camera-l515/ [Accessed: June 25, 2020].

Kang, J. W. & Liu, B. (2014). 'Application of 3D stereoscopic visualization technology in casting aspect' [Online]. *China Foundry*. Available at: https://www.researchgate.net/publication/287049069_Application_of_3D_stereoscopic_visualization_technology_in_casting_aspect [Accessed: 4 July 2020].

Kim, D. & Lee, S. (2013). 'Advances in 3D Camera: Time-of-Flight vs. Active Triangulation', *Intelligent Autonomous Systems 12: Volume 1 Proceedings of the 12th International Conference IAS-12, 26-29 June 2012, Jeju Island, Korea*. [Online]. Berlin: Springer. pp. 301–309. Available at: doi:10.1007/978-3-642-33926-4_28 [Accessed: 2 July 2020].

MacRae, I. & Murphy, T. A. (2017). Does employee surveillance improve or impair performance? *Personnel Today* [Online]. Available at: https://www.personneltoday.com/hr/employee-surveillance-improve-impair-performance/ [Accessed: 25 June 2020].

Manifold, S. (2020). *SmartSpace* [Online]. Available at: https://ubisense.com/smartspace/ [Accessed: 23 June 2020].

Markets and Markets (2019). 'Smart Space Market by Component (Solutions and Services), Application (Energy Optimization and Management, Emergency Management, and Security Management), Premises Type (Commercial, Residential), and Region - Global Forecast to 2024' [Online]. Available at: https://www.marketsandmarkets.com/Market-Reports/smart-space-market-252626748.html [Accessed: 9 June 2020].

Microsoft (2020). *Azure Kinect DK hardware specifications* [Online]. Available at: https://docs.microsoft.com/en-us/azure/kinect-dk/hardware-specification [Accessed: 25 June 2020].

Mordor Intelligence (2019). 'Smart Space Market Size, Share, Trends, Forecast (2019-24)' [Online]. Available at: https://www.mordorintelligence.com/industry-reports/smart-space-market [Accessed: 21 June 2020].

MoviMed (2018). *What is Structured Light Imaging?* [Online]. Available at: https://roboticstomorrow.com/article/2018/04/what-is-structured-light-imaging/11821 [Accessed: 3 July 2020].

Murthy, C. B., Hashmi, M. F., Bokde, N. D., & Geem, Z. W. (2020). 'Investigations of Object Detection in Images/Videos Using Various Deep Learning Techniques and Embedded Platforms – A comprehensive Review', *Applied Sciences*, 10 (9), https://doi.org/10.3390/app10093280 [Accessed: 14 August 2020].

Myers, B. (2018). 'Capturing depth: structured light, time of flight, and the future of 3D imaging' [Online]. Android Authority. 22 June. Available at: https://www.androidauthority.com/structured-light-3d-imaging-870016/ [Accessed: 1 July 2020].

Mynt Eye (n.a.). *MYNT EYE D Depth Camera* [Online]. Available at: https://www.mynteye.com/pages/mynt-eye-d [Accessed: 25 June 2020].

O' Riordan, A., Newe, T., Toal, D. & Dooly, G. (2018). 'Stereo Vision Sensing: Review of existing systems', *Twelfth International Conference on Sensing Technology (ICST)*, University of Limerick, 4 December 2018. Ollscoil Luimnigh. Available at: doi:10.1109/ICSensT.2018.8603605.

Orbbec (n.a.). *Astra Series* [Online]. Available at: https://orbbec3d.com/product-astra-pro/ [Accessed: 25 June 2020].

Parletta, N. (2019). 'Going Green - What's Good For The Planet Is Good For Business'. *Forbes*. 14 May. Available at: https://www.forbes.com/sites/natalieparletta/2019/05/14/going-green-whats-good-for-the-planet-is-good-for-business/ [Accessed: 15 June 2020].

Porter, M. E. & Heppelmann, J. E. (2014). How Smart, Connected Products Are Transforming Competition. *Harvard Business Review* [Online]. Available at: https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition [Accessed: 15 June 2020].

Rake, R. & Jagtap, B. (2019). 'Smart Space Market Size, Share and Industry Analysis | Forecast – 2026' [Online]. Allied Market Research. Available at: https://www.alliedmarketresearch.com/smart-space-market [Accessed: 15 June 2020].

ROS-Industrial (2020). '3D Camera Survey' [Online]. Available at: https://rosindustrial.org/3d-camera-survey [Accessed: 25 June 2020].

Roth, C. (2018). 'How to market and sell digital workplace solutions'. *Reseller News*, 21 May. Available at: https://www.reseller.co.nz/article/641351/how-market-sell-digital-workplace-solutionshow-market-sell-digital-workplace-solutions/ [Accessed: 15 June 2020].

Schneider Electric (2019). *Financial and Sustainable Development Report 2019* [Online]. Available at: https://annualreport.se.com/#download-centre [Accessed: 21 June 2020].

Schneider Electric (2020). *EcoStruxure: IoT – Internet of Things* [Online]. Available at: https://www.se.com/ww/en/work/campaign/innovation/overview [Accessed: 21 June 2020].

Sharma, B. (2019). 'What is LiDAR technology and how does it work?' [Online]. *Geospatial World*. 30 January. Available at: https://www.geospatialworld.net/blogs/what-is-lidar-technology-and-how-does-it-work/ [Accessed: 2 July 2020].

Siemens (2019). *Annual Report 2019* [Online]. Available at: https://assets.new.siemens.com/siemens/assets/api/uuid:59a922d1-eca0-4e23-adef-64a05f0a8a61/siemens-ar2019.pdf [Accessed: 21 June 2020].

Siemens Advanta (2020). *About Us* [Online]. Available at: https://www.siemens-advanta.com/ [Accessed: 21 June 2020].

SmartSpace Software PLC (2019). *Innovative Workplace Solutions - SmartSpace Software Plc* [Online]. Available at: https://www.smartspaceplc.com [Accessed: 23 June 2020].

Spacewell International (2020). *Home* [Online]. Available at: https://spacewell.com/ [Accessed: 23 June 2020].

StereoLabs (2020). *ZED 2 Camera and SDK Overview* [Online]. Available at: https://cdn.stereolabs.com/assets/datasheets/zed2-camera-datasheet.pdf [Accessed: June 25, 2020].

Tillman, M. (2020). 'What is a ToF camera? The Time-of-flight sensor explained' [Online]. *Pocket-lint*. 3 August. Available at: https://www.pocket-lint.com/phones/news/147024-what-is-a-time-of-flight-camera-and-which-phones-have-it [Accessed: 1 July 2020].

Vergesense (2020). *Solutions* [Online]. Available at: https://vergesense.com/solutions [Accessed: 23 June 2020].

Verschae, R. & Ruiz-del-Solar, J. (2015). 'Object Detection: Current and Future Directions', *Frontiers in Robotics and AI*, 2, pp. 29, https://doi.org/10.3389/frobt.2015.00029 [Accessed: 14 August 2020].

Wellener, P., Michalik, J, Manolian, H. J. & Griffin, J. (2018). 'Smart Buildings | Deloitte Insights.' [Online]. Available at: https://www2.deloitte.com/content/dam/Deloitte/br/Documents/financial-services/DI_Smart-buildings.pdf [Accessed: 9 June 2020].

Yanzi (2020). *Yanzi Platform* [Online]. Available at: https://www.yanzi.se/ [Accessed: 23 June 2020].

Sources included in the Systematic Literature Review:

Aksenov, A.Y., Kuleshov, S.V. & Zaytseva, A.A. (2014). 'An Application of Computer Vision Systems to Solve the Problem of Unmanned Aerial Vehicle Control', *Transport and Telecommunication,* 15 (3), pp. 209-214.

Becker, B.C. (2012). *Vision-Based Control of a Handheld Micromanipulator for Robot-Assisted Retinal Surgery*. PhD Thesis. Carnegie Mellon University (USA).

Bharatharaj, J., Huang, L., Mohan, R.E., Pathmakumar, T., Krägeloh, C. & Al-Jumaily, A. (2018). 'Head Pose Detection for a Wearable Parrot-Inspired Robot Based on Deep Learning', *Applied Sciences, 8* (7)

Bleser, G., Damen, D., Behera, A., Hendeby, G., Mura, K., Miezal, M., Gee, A., Petersen, N., Maçães, G., Domingues, H., Gorecky, D., Almeida, L., Mayol-Cuevas, W., Calway, A., Cohn, A.G., Hogg, D.C. & Stricker, D. (2015). 'Cognitive Learning, Monitoring and Assistance of Industrial Workflows Using Egocentric Sensor Networks', *PLoS One,* 10 (6).

Cai, M., Lu, F. & Gao, Y. (2019). 'Desktop Action Recognition From First-Person Point-of-View', *IEEE Transactions on Cybernetics*, 49 (5), pp. 1616-1628, doi: 10.1109/TCYB.2018.2806381.

Carneiro, D., Novais, P., Augusto, J. C. & Payne, N. (2019). 'New Methods for Stress Assessment and Monitoring at the Workplace', IEEE *Transactions on Affective Computing,* 10 (2), pp. 237-254, doi: 10.1109/TAFFC.2017.2699633.

Choi, J. W., Quan X. & Cho, S. H. (2018). 'Bi-Directional Passing People Counting System Based on IR-UWB Radar Sensors', *IEEE Internet of Things Journal*, 5 (2), pp. 512-522, doi: 10.1109/JIOT.2017.2714181.

Cipriani, A., & Geddes, J. (2003). 'Comparison of systematic and narrative reviews: The example of the atypical antipsychotics'. *Epidemiologia E Psichiatria Sociale*, 12(3), 146-153.

DattaGupta, A. (2012). *Stereo matching - Improving image quality*. PhD Thesis. The University of Texas at El Paso (USA).

Edirisinghe, R. (2019). 'Digital skin of the construction site', *Engineering, Construction and Architectural Management*, *27*(2), pp. 184-223.

Faccio, M., Ferrari, E., Gamberi, M. & Pilati, F. (2019). 'Human Factor Analyser for work measurement of manual manufacturing and assembly processes', *The International Journal of Advanced Manufacturing Technology,* 103 (1-4), pp. 861-877.

Glaser, B. & Strauss, A. (1967). *The Discovery of Grounded Theory* (reprinted 2006). New Brunswick: AldineTransaction.

Grant, M. J. & Booth, A. (2009). 'A typology of review: an analysis of 14 review types and associated methodologies' [Online]. *Health Information and Libraries*, 2 (26), https://onlinelibrary.wiley.com/doi/full/10.1111/j.1471-1842.2009.00848.x

Hogreve, S., Kaczmarek, S., Adam, J., Franz, L., Döllen, T., Paulus, H., Reinkemeyer, V. & Tracht, K. (2016). ‘Controlling and Assisting Manual Assembly Processes by Automated Progress and Gesture Recognition’, *Applied Mechanics and Materials,* 840, pp. 50-57.

Jackson, R. C., Yuan, R., Chow, D., Newman, W. S. & Çavuşoğlu, M. C. (2018). ‘Real-Time Visual Tracking of Dynamic Surgical Suture Threads’, *IEEE Transactions on Automation Science and Engineering*, 15 (3), pp. 1078-1090, doi: 10.1109/TASE.2017.2726689.

Jacobsen, M.D. (2014). *Smart Frame Grabber: A Hardware Accelerated Computer Vision Framework*. PhD Thesis. University of California, San Diego (USA).

Kaczmarek, A.L. (2020). ‘3D Vision System for a Robotic Arm Based on Equal Baseline Camera Array’, *Journal of Intelligent & Robotic Systems,* 99(1), pp. 13-28.

Kasaei, S.H.M. (2019). *Interactive Open-Ended Learning for 3D Object Recognition.* PhD Thesis. Universidade de Aveiro (Portugal).

Kim, K., Kim, J., Choi, J., Kim, J. & Lee, S. (2015). ‘Depth Camera-Based 3D Hand Gesture Controls with Immersive Tactile Feedback for Natural Mid-Air Gesture Interactions’, *Sensors,* 15 (1), pp. 1022-1046.

Kouskouridas, R., Gasteratos, A. & Badekas, E. (2012). ‘Evaluation of two-part algorithms for objects' depth estimation’, IET *Computer Vision*, 6 (1), pp. 70-78, doi: 10.1049/iet-cvi.2009.0094.

Krejov, P., Gilbert, A. & Bowden, R. (2014). ‘A Multitouchless Interface: Expanding User Interaction’, *IEEE Computer Graphics and Applications*, 34 (3), pp. 40-48, doi: 10.1109/MCG.2014.44.

Kyrkou, C., Christoforou, E. G., Timotheou, S., Theocharides, T., Panayiotou, C. & Polycarpou, M. (2018). ‘Optimizing the Detection Performance of Smart Camera Networks Through a Probabilistic Image-Based Model’, *IEEE Transactions on Circuits and Systems for Video Technology*, 28 (5), pp. 1197-1211, doi: 10.1109/TCSVT.2017.2651362.

Lai, K.K.W. (2013). *Object Recognition and Semantic Scene Labeling for RGB-D Data*. PhD Thesis. University of Washington (USA).

Liu, S. (2019). ‘Computer Vision for Perception and Localization’, *Engineering Autonomous Vehicles and Robots: The DragonFly Modular-based Approach*, pp.77-96, doi: 10.1002/9781119570516.ch6.

Liu, T. (2018). *An Adaptive Human Intention Recognition Framework to Facilitate Human-robot Cooperative Assembly Tasks.* PhD Thesis. The Chinese University of Hong Kong (Hong Kong).

Lun, R.Z. (2017). *Human Activity Tracking and Recognition Using Kinect Sensor*. PhD Thesis. Cleveland State University (USA).

Maereg, A.T., Secco, E.L., Agidew, T.F., Reid, D. & Nagar, A.K. (2017). ‘A Low-Cost, Wearable Opto-Inertial 6-DOF Hand Pose Tracking System for VR’, *Technologies, 5* (3), pp. 49.

Maki, J.N., Golombek, M., Deen, R., Abarca, H., Sorice, C., Goodsall, T., Schwochert, M., Lemmon, M., Trebi-Ollennu, A. & Banerdt, W.B. (2018). ‘The Color Cameras on the InSight Lander’, *Space Science Reviews, 214* (6), pp. 1-34.

Mbouzao, B. (2013). *Quantitative Assessment of Human Motion Capabilities with Passive Vision Monitoring*. PhD Thesis. University of Ottawa (Canada).

Mei, C., Sibley, G., Cummins, M., Newman, P. & Reid, I. (2011). ‘RSLAM: A System for Large-Scale Mapping in Constant-Time Using Stereo’, *International Journal of Computer Vision, 94* (2), pp. 198-214.

Mendez, O. (2017). *Collaborative Strategies for Autonomous Localisation, 3D Reconstruction and Pathplanning*. PhD Thesis. University of Surrey (United Kingdom).

Narayana, P. (2018). *Improving Gesture Recognition Through Spatial Focus of Attention*. PhD Thesis. Colorado State University (USA).

Opoku, D. (2013). *Novel Approach to Intelligent Navigation of a Mobile Robot in a Dynamic and Cluttered Indoor Environment*. PhD Thesis North Carolina Agricultural and Technical State University (USA).

Oszust, M., Kapuscinski, T., Warchol, D., Wysocki, M., Rogalski, T., Pieniazek, J., Kopecki, G.H., Ciecinski, P. & Rzucidlo, P. (2018). 'A vision-based method for supporting autonomous aircraft landing', *Aircraft Engineering and Aerospace Technology, 90* (6), pp. 973-982.

Paleja, R.R. (2018). *Viability and Performance of Indoor Mapping using the Velodyne VLP-16 LiDAR*. PhD Thesis. Rutgers The State University of New Jersey, School of Graduate Studies (USA).

Patil, V. & Khan, I.A. (2016). 'HAND GESTURE RECOGNITION', *BVIMSR's Journal of Management Research, 8* (1), pp. 76-82.

Perez, L., Rodríguez, Í., Rodríguez, N., Usamentiaga, R. & García, D., F. (2016). 'Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review', *Sensors,* 16 (3), pp. 335.

Pradhan, N. (2013). *Mobile Robot Navigation for Person Following in Indoor Environments*. PhD Thesis. Clemson University (USA).

Rafibakhsh, N. (2013). *Industrial Applications of Microsoft Xbox Kinect Sensor*. PhD Thesis. Southern Illinois University at Edwardsville (USA).

Raza, R.H. (2013). *Three-dimensional localization and tracking for site safety using fusion of computer vision and RFID*. PhD Thesis. Michigan State University (USA).

Reiter, A. (2013). *Assistive Visual Tools for Surgery*. PhD Thesis. Columbia University (USA).

Roudaki, A. (2013). *Spatially aware computing for natural interaction*. PhD Thesis. North Dakota State University (USA).

Roy, P. (2019). *Computer Vision Algorithms for Yield Mapping in Specialty Farms*. PhD Thesis. University of Minnesota (USA).

Scano, A., Robert, M.M., Cerveri, P., Lorenzo, M.T. & Sacco, M. (2020). 'Analysis of Upper-Limb and Trunk Kinematic Variability: Accuracy and Reliability of an RGB-D Sensor'. *Multimodal Technologies and Interaction, 4* (2), pp. 14.

Schulze, M., Nehler, H., Ottosson, M. & Thollander, P. (2016). 'Energy management in industry – a systematic review of previous findings and an integrative conceptual framework'. *Journal of Cleaner Production,* 112, pp. 3692–3708. Available at: doi:10.1016/j.jclepro.2015.06.060

Seo, J. (2016). *Evaluation of construction workers' physical demands through computer vision-based kinematic data collection and analysis*. PhD Thesis. University of Michigan (USA).

Shim, J., Yang, Y., Kang, N., Seo, J. & Han, T. (2016). 'Gesture-based interactive augmented reality content authoring system using HMD', *Virtual Reality, 20* (1), pp. 57-69.

Sikdar, A. (2018). *Depth based Sensor Fusion in Object Detection and Tracking*. PhD Thesis. The Ohio State University (USA).

Silva, A.E. (2014). *Collaborative human-machine quality control system: Steps towards automatic machine vision inspection*. PhD Thesis. Universidade do Porto (Portugal).

Sun, Y., Hollerbach, J. M. & Mascaro, S. A. (2019). ‚Estimation of Fingertip Force Direction With Computer Vision', *IEEE Transactions on Robotics*, 25 (6), pp. 1356-1369, doi: 10.1109/TRO.2009.2032954.

Wang, J. (2016). *Novel Strategies for Enhanced Navigation in Robotic Surgical System*. PhD Thesis. The Chinese University of Hong Kong (Hong Kong).

Wang, J., Song, S., Ren, H., Lim, C. M. & Meng, M. Q. (2019). 'Surgical Instrument Tracking By Multiple Monocular Modules and a Sensor Fusion Approach', *IEEE Transactions on Automation Science and Engineering*, 16 (2), pp. 629-639, doi: 10.1109/TASE.2018.2848239.

Xiang, T., Xia, G. & Zhang, L. (2019). 'ini-Unmanned Aerial Vehicle-Based Remote Sensing: Techniques, applications, and prospects', *IEEE Geoscience and Remote Sensing Magazine*, 7 (3), pp. 29-63, doi: 10.1109/MGRS.2019.2918840.

Xing, Y. *et al*. (2018). 'Advances in Vision-Based Lane Detection: Algorithms, Integration, Assessment, and Perspectives on ACP-Based Parallel Vision', *IEEE/CAA Journal of Automatica Sinica*, 5 (3), pp. 645-661, doi: 10.1109/JAS.2018.7511063.

Yamafune, K. (2016). *Using computer vision photogrammetry (Agisoft PhotoScan) to record and analyze underwater shipwreck sites*. PhD Thesis. Texas A&M University (USA).

Yang, G., Lv, H., Chen, F., Pang, Z., Wang, J., Yang, H. & Zhang, J. (2018). 'A Novel Gesture Recognition System for Intelligent Interaction with a Nursing-Care Assistant Robot', *Applied Sciences, 8* (12)

Zhang, X. (2015). *Facial action understanding from human perception and machine vision perspectives*. PhD Thesis. State University of New York at Binghamton (USA).

Appendices

Appendix A: Methodology of the Competitive Positioning Map

The Competitive Positioning Map consists of a dot-diagram partitioned in four groups (Outpacer, funcationality leader, cost leader, and flop). Its x-axis resembles the selling price of the device and its y-axis reflects the weighted functionality score. The selling price does not require further explanation other then it was recorded in June 2020 from the manufacturer's website. The weighted functionality score needs some further explanation, how it was derived.

The functionality score was created out of a weighted point scoring system, which in term is dependent on a ranking. The ranking was established for the following categories: RGB Video Camera, Frame Rate (of the Video Camera), Depth-Range, 3D-Resolution, Frame Rate (of the Depth-Sensor), 3D-Resolution, Microphone, Accelerometer, Gyroscope, Other Sensors, SDK, Physical Dimensions, and Weight. All those categories provided clear technical data about their functionality. As there was no technical data about the accelerometer, gyroscope, and other integrated sensors available for all devices, those categories were judged in a binary method. Devices received functionality points if they have those sensors or none if they do not possess those sensors. This minimized any researcher bias. The ranking itself was established according to standard statistical methods laid out according to the Center for Applied Statistics Courses UCL (2018). The table below depicts the ranking across the categories.

Table 2. Ranking of sensor sample.

Sensor Name	Azure Kinect DK	ASUS XtionPro Live	ZED 2 Stereo Camera	e-Con Systems Tara Stereo Camera	Orbbec Astra Series	Intel RealSense L515	Intel RealSense D435i/ D435	Mynt Eye D-1000-120 FOV
RGB Video Camera	1	6	2	8	7	3,5	3,5	5
Frame Rate	7	5	1	8	5	3	5	2
Depth-Sensor								
Type								
Depth-Range	6	7	1	8	5	4	3	2
3D-Resolutions	2	7,5	1	6	7,5	5	3,5	3,5
Frame Rate	6,5	6,5	1	3,5	6,5	6,5	2	3,5
Microphone	1	2,5	6	6	2,5	6	6	6
Accelerometer	1	0	1	0	0	1	1	1
Gyroscope	1	0	1	0	0	1	1	1
Other Sensors	0	0	1	0	0	0	0	0
SDK	1	7,5	4,5	7,5	6	2,5	2,5	4,5
Physical Dimensions	8	7	5	3	6	1	2	4
Weight	7	8	4	2	6	3	1	5

Next, the ranking was translated into a point scoring system. Below, the first table maps the ranking to a score. The second table applies the scoring system to the ranking table and depicts the final result.

Table 3. Scoring system mapping a rank to a score.

Ranking	Score
1	8
1,5	7,5
2	7
2,5	6,5
3	6
3,5	5,5
4	5
4,5	4,5
5	4
5,5	3,5
6	3
6,5	2,5
7	2
7,5	1,5
8	1

Table 4. Scores of the sensor sample.

Sensor Name	Azure Kinect DK	ASUS XtionPro Live	ZED 2 Stereo Camera	e-Con Systems Tara Stereo Camera	Orbbec Astra Series	Intel RealSense L515	Intel RealSense D435i/ D435	Mynt Eye D-1000-120 FOV
RGB Video Camera	8	3	7	1	2	5,5	5,5	4
Frame Rate	2	4	8	1	4	6	4	7
Depth-Sensor								
Type								
Depth-Range	3	2	8	1	4	5	6	7
3D-Resolutions	7	1,5	8	3	1,5	4	5,5	5,5
Frame Rate	2,5	2,5	8	5,5	2,5	2,5	7	5,5
Microphone	8	6,5	3	3	6,5	3	3	3
Accelerometer	8	0	8	0	0	8	8	8
Gyroscope	8	0	8	0	0	0	8	8
Other Sensors	0	0	8	0	0	8	0	0
SDK	8	1,5	4,5	1,5	3	6,5	6,5	4,5
Physical Dimensions	1	2	4	6	3	8	7	5
Weight	2	1	5	7	3	6	8	4

In a last step, a weighted score was calculated for each device. The first table underneath contains the weights for each category. The second table shows the weighted functionality scores for each of the sample devices, plotted in the Competitive Positioning Map.

Table 5. Weighting of the categories.

Category	Weighting
RGB Video Camera	0,2
Frame Rate	0,05
Depth-Sensor	n.a.
Type	n.a.
Depth-Range	0,15
3D-Resolutions	0,2
Frame Rate	0,05
Microphone	0,05
Accelerometer	0,05
Gyroscope	0,05
Other Sensors	0,05
SDK	0,1
Physical Dimensions	0,025
Weight	0,025

Table 6. Data collection for the Competitive Positioning Map.

Sensor Name	Sum of Functionality Score (out of 96)	Sum of Weighted Functionality Score (max. 8)	Weighted Functionality Score Rank	Price
Azure Kinect DK	57,5	5,75	3	$ 399
ASUS XtionPro Live	24	2,075	7	$ 300
ZED 2 Stereo Camera	79,5	7,025	1	$ 449
e-Con Systems Tara Stereo Camera	29	1,9	8	$ 149
Orbbec Astra Series	29,5	2,4	6	$ 200
Intel RealSense L515	62,5	5,025	5	$ 349
Intel RealSense D435i/D435	68,5	5,625	2	$ 199
Mynt Eye D-1000-120 FOV	61,5	5,2	4	$ 389

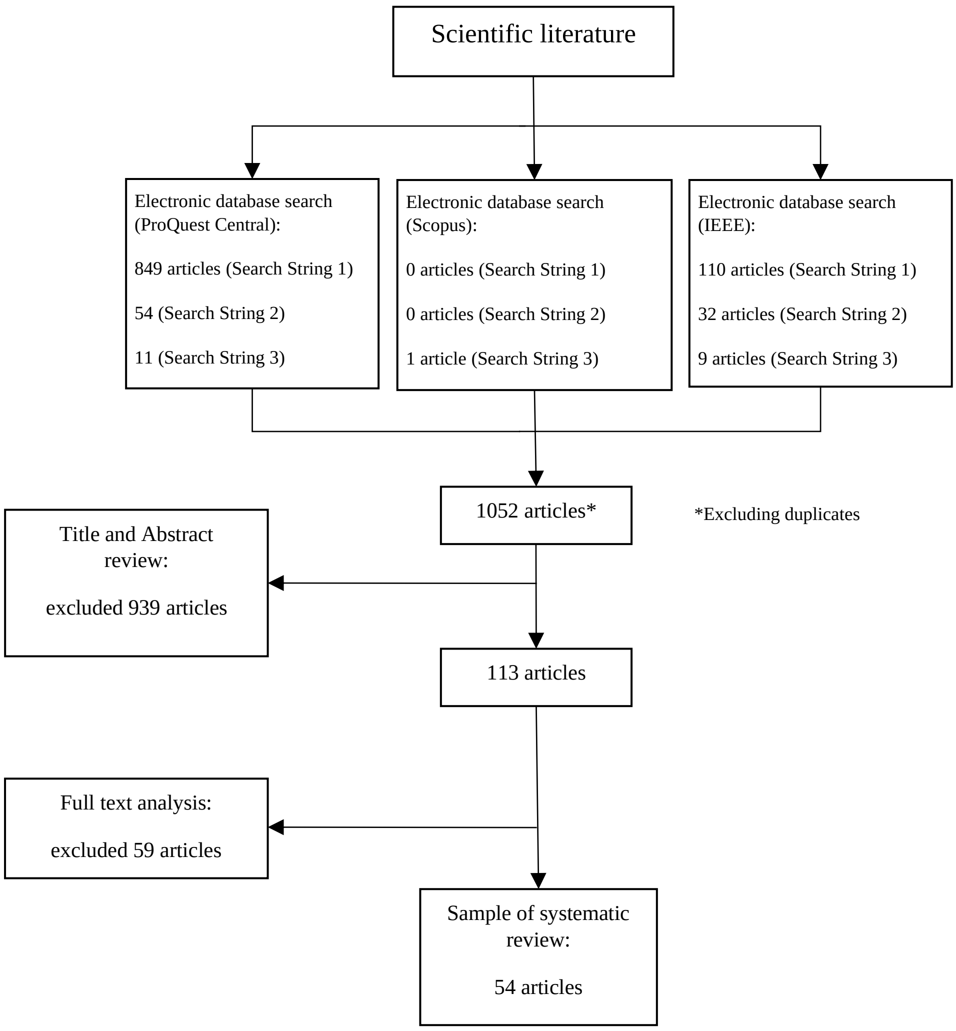
Appendix B: Search Strings for the Literature Review

Search String 1	"computer vision" AND workspace AND usage	Searched in Title, Abstract, Keywords, and Full Text
Search String 2	"computer vision" AND "physical space" AND utilization	Searched in Title, Abstract, Keywords, and Full Text
Search String 3	"computer vision" AND "physical space" AND "sensor technology"	Searched in Title, Abstract, Keywords, and Full Text

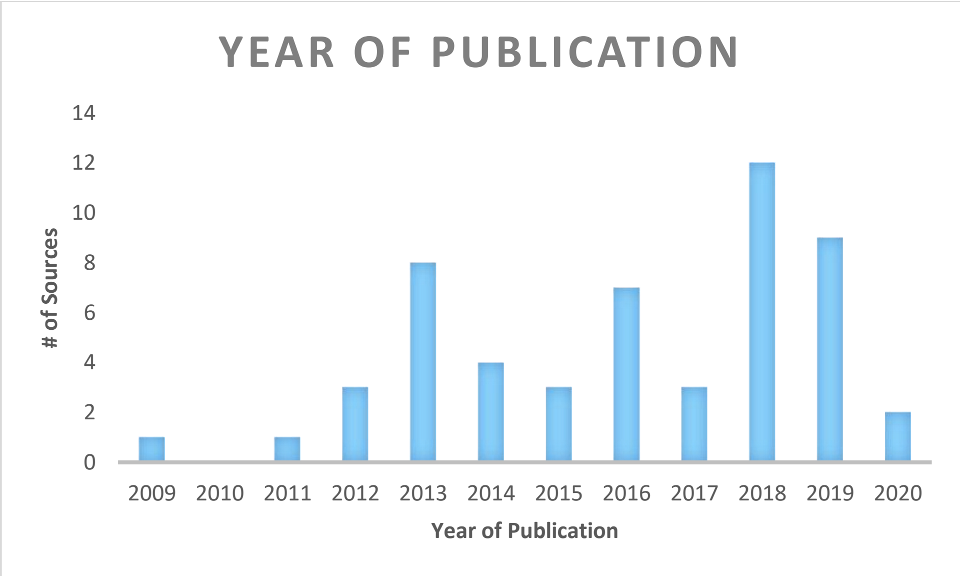
Appendix C: Overview of Inclusion and Exclusion Criteria of the Literature Review

	Inclusion	**Exclusion**
Publication Type	Peer-reviewed academic journal, working papers, dissertations, scholarly magazines/books	Any kind of other publications (e.g. periodicals, citations, new wires)
Language	English	All other languages
Accessibility / Availability	Online available as full text	If not available in full text* *Note: If only sections of the text were available and these had a sufficient level of detail for the investigation, then they were included.
Research Discipline	No specification	-
Time Period	2009 – 2020* *Note: The thesis was written between June and September 2020. Everything up to this time period was included.	-
Sector	No specification	-
Relevance	• Resource discusses at minimum technological developments in the field of 3D/Depth sensing technology, combined with a use case • Resource addressed novel use cases applying existing depth sensing technology	• Resource solely discusses theoretical concepts of future developments in the field 3D/Depth sensing technology • Resource did not mention any possible use cases or improvements of existing computer vision algorithms

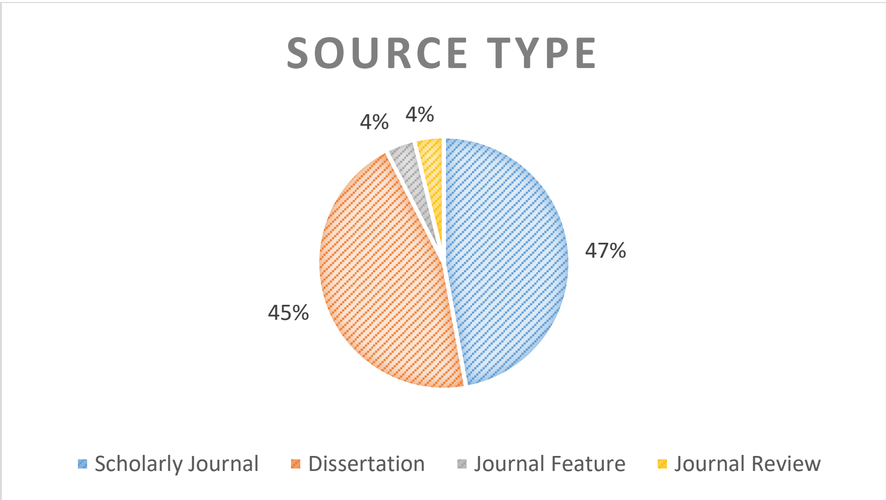
Appendix D: Overview of the Search Strategy of the Literature Review (figure structure adapted from Schulze et al. (2015))



Appendix E: Graph Displays Time of Publication of the Literature Sample



Appendix F: Pie Chart Displaying Source Type of the Sample Literature



Appendix F: Table Matching Identified Use Cases to Sources of the Literature Review

In Literature Review Identified Use Cases for Computer Vision Technology	**Source**
Agriculture	(Roy, 2019)
Archaeology	(Yamafune, 2016)
Autonomous Aerial Vehicles	(Xiang, Xia & Zhang, 2019), (Oszust et al., 2018), (Aksenov, Kuleshov & Zaytseva, 2014)
Autonomous Driving	(Xing et al., 2018)
Real Estate/Building	(Cai, Lu & Gao, 2019), (Choi, Quan & Cho, 2018)
Computer Vision Application Development	(Jacobsen, 2014), (Lai, 2013), (DattaGupta, 2012), (Sun, Hollerbach & Mascaro, 2009), (Kouskouridas, Gasteratos & Badekas, 2012), (Kyrkou et al., 2018)
Construction Site	(Edirishinghe, 2019), (Seo, 2016)
Education	(Patil & Khan, 2016)
Human-Computer-Interaction	(Kwangtaek et al., 2015), (Roudaki, 2013), (Zhang, 2015), (Krejov, Gilbert & Bowden, 2014)
Human-Robot-Interaction	(Bharatharaj et al., 2018), (Liu, 2018), (Narayana, 2018), (Geng et al., 2018)
Indoor Mapping	(Paleja, 2018)
(Industrial) Manufacturing	(Hogreve et al., 2016), (Bleser et al., 2015), (Rafibakhsh, 2013), (Faccio, Ferrari, Gamberi & Pilati, 2019), (Silva, 2014)
Medicine	(Scano et al., 2020), (Mbouzao, 2013), (Wang, et al., 2019), (Jackson et al., 2018), (Wang, 2016), (Reiter, 2013), (Becker, 2012)
Outdoor Mapping	(Mei et al., 2011)
Perceptual Learning	(Kasaei, 2019)
Robotic Vision	(Kaczmarek, 2020), (Liu, 2019), (Maki et al., 2018), (Sikdar, 2018), (Mendez, 2017), (Perez et al., 2016), (Pradhan, 2013), (Opoku, 2013)
Surveillance	(Raza, 2013)
VR/AR	(Maereg, 2017), (Jinwook et al., 2016)
Worker Wellbeing	(Lun, 2017), (Carneiro, Novais, Augusto & Payne, 2019)

Appendix G: Table Matching Identified Computer Vision Technologies to Sources of the Literature Review

In Literature Review Identified Computer Vision Technology Groups	Source
Body Tracking	(Scano et al., 2020), (Cai, Lu & Gao, 2019), (Bharatharaj et al., 2018), (Liu, 2018), (Lun, 2017), (Maereg, 2017), (Seo, 2016), (Mbouzao, 2013)
Computer Vision Algorithm	(Kaczmarek, 2020), (Kyrkou et al., 2018), (Aksenov, Kuleshov & Zaytseva, 2014), (Jacobsen, 2014), (Pradhan, 2013), (Becker, 2012), (DattaGupta, 2012), (Kouskouridas, Gasteratos & Badekas, 2012)
Gesture Recognition	(Carneiro, Novais, Augusto & Payne, 2019), (Narayana, 2018), (Geng et al., 2018), (Jinwook et al., 2016), (Kwangtaek et al., 2015), (Patil & Khan, 2016), (Zhang, 2015), (Krejov, Gilbert & Bowden, 2014), (Sun, Hollerbach & Mascaro, 2009)
Mapping	(Paleja, 2018), (Mei et al., 2011)
Object Recognition	(Kasaei, 2019), (Roy, 2019), (Sikdar, 2018), (Xing et al., 2018), (Choi, Quan & Cho, 2018)
Object Scanning	(Yamafune, 2016)
Object Tracking	(Wang, et al., 2019), (Jackson et al., 2018), (Oszust et al., 2018), (Wang, 2016), (Raza, 2013), (Reiter, 2013)
Process Tracking	(Edirishinghe, 2019), (Hogreve et al., 2016), (Bleser et al., 2015), (Rafibakhsh, 2013), (Faccio, Ferrari, Gamberi & Pilati, 2019)
Quality Control	(Silva, 2014)
Sensor Technology	(Maki et al., 2018), (Perez et al., 2016), (Lai, 2013)
Spatial Perception	(Liu, 2019), (Xiang, Xia & Zhang, 2019), (Mendez, 2017), (Roudaki, 2013), (Opoku, 2013)