University of Twente

BSc Thesis Biomedische Technologie

Remote Real-Time 3D Viewing During Surgery for Supervision and Education

Applications in Endoscopy and Remote Viewing via Camera-equipped Glasses

R.M.M. Tol, s2539594

Examination Committee
Prof. dr. ir. R.M. Verdaasdonk
Rob van Doremalen
Sigert Mevissen

July 4, 2024

Abstract

This thesis explores the implementation of remote real-time 3D viewing technology for supervision and education applications in endoscopy and remote viewing via camera-equipped glasses. Stereoscopic imaging is important in surgeries where precision is required. This project builds on existing technologies for using 3D in operating rooms, such as remote viewing via an operating microscope at Deventer Hospital.

Humans are able to see 3D because each eye sees a different part of the world. The brain combines these two images into a single image in which depth can be perceived. This is called stereoscopy. Using two cameras, the surgical field is recorded from different angles to get a stereoscopic image. In this project, this is done with an endoscope and camera-equipped glasses. The two images are brought in via software programs on the laptop and pasted next to each other to get a side-by-side view. This image is then streamed to a phone placed in a VR headset. This allows live viewing during operations to give advise or learn from it.

The thesis covers the technical implementation, required hardware and software and the additional challenges, both technical, ethical and legal. Technical challenges include latency, image quality maintenance and camera alignment. Ethical and legal challenges include compliance with various laws and regulations that need to be adhered to. These include the GDPR to protect personal data and the MDR to ensure the safety of medical devices.

In addition, tests were done to validate the performance of the systems. Test rigs were built for both the endoscope and the camera-equipped glasses. Resolution, colour accuracy, depth of field, image distortion, latency and depth perception were tested according to predetermined program requirements. The results showed that the glasses scored well on image distortion and latency, but less on resolution, colour accuracy and depth perception. The endoscope performed well on resolution, depth of field, image distortion and depth perception, but lower on colour and latency.

It also places the results in clinical context and suggests possible points for improvement and future perspectives including possible follow-up studies.

Contents

List of Figures

List of Tables

1 Introduction

This project focuses on developing and optimising stereoscopic imaging techniques for real-time remote viewing during surgical procedures, with applications in endoscopy, surgical microscopes and camera-equipped glasses.

Stereoscopic imaging is important in various types of surgery, such as ear, nose and throat (ENT) surgery, microsurgery and endoscopy. These forms of surgery require precision to repair or view very small or complex structures. Stereoscopic imaging in hospitals is created by using a surgical microscope or endoscope that can capture images in 3D and display them on a monitor. This ensures that surgeons have good depth perception, which allows them to work very precisely. This reduces the chances of mistakes that can occur, leads to fewer complications and facilitates a faster recovery.

At Deventer Hospital, VR goggles are already being used to allow ENT specialists to view ear surgeries live remotely. This enables live 3D images to be shown from the surgical microscope, improving cooperation and communication between doctors. The technology is currently still in the testing phase, but doctors are excited about the sharpness and depth of the images.
The use of 3D in healthcare is not a new phenomenon. In operating rooms (OR), monitors are used on which OR staff and medical students can see the same depth as the operating surgeon using 3D glasses. The difference is that live 3D images can be streamed to VR goggles via a phone connection.
Before this technology can be widely used, it has to be medically approved, which could take several years. Ultimately, this technology could represent a significant advancement for healthcare and medical education, especially in complex surgeries where external expertise is needed, such as eye and plastic surgery. [1]
This project builds on these existing technologies by implementing real-time 3D viewing in endoscopy and remote viewing via camera-equipped glasses.

The use of real-time 3D supervision enables surgeons to obtain a comprehensive and accurate view of the surgical field. This allows them to better visualise complex structures, perceive anatomical details and identify potential complications in a timely manner [2]. Consequently, this leads to improved decision-making during the operation, which in turn can enhance surgical outcomes such as faster patient recovery and reduced complications [3].
Real-time 3D viewing also allows external observers, such as other surgeons, medical students, and researchers to remotely follow the surgery and directly observe the techniques and procedures being performed. This expands opportunities for education, mentorship and knowledge transfer within the medical community.
So, complex surgical procedures can be documented and shared for educational purposes, enhancing the learning experience and raising the level of surgical expertise [4, 5].

The project involves adaptation of stereoscopic imaging systems using two cameras to capture the surgical field from different angles like left and right eye view. This technique will be integrated into endoscopic devices and wearable glasses with cameras, enabling the acquisition of a stereoscopic video stream in real-time. This stream can be viewed on a smart phone placed inside a virtual reality (VR) goggles, providing a stereoscopic view of the surgical field in real-time.

Streaming live 3D images involves several challenges, including technical and regulatory aspects. Technical challenges include image quality and latency, which are crucial for realistic representation of the surgical field. With regulatory issues, it is important to ensure security and privacy so that the confidentiality of medical data is not compromised.

1.1 Problem definition

In certain cases, access to performing complex surgeries is limited, especially in rural areas or in low-resource countries. Remote live streaming can provide a solution to this challenge: surgeons can be supported or advised during operations without the need for the supervisor or specialised surgeon to be physically present. Also, live streaming has been found to be more beneficial for student learning than watching a video afterwards. Being able to view the surgery in real-time allows students to better absorb and comprehend the surgical procedures and techniques being performed. Live-streaming can also reduce variability in students' surgical experiences by providing a consistent and real-time learning environment. [6]

Equipment for capturing 3D recordings is often bulky and cumbersome. This is a significant disadvantage for surgeons as it hinders their work. Moreover, they often encounter complex software that disrupts the natural workflow in the operating room. Existing solutions are often incompatible with headlights, masks and face shields, which further obstruct surgeons' movements and visibility. Ideally, a system should be simple to use for both the surgeon and observers. [5, 6]

In operating rooms, it can be challenging to accommodate everyone who wants to observe, especially since these rooms quickly become crowded [7]. However, with cameras, surgeons and others involved no longer need to be present in the operating room to follow the procedure. This facilitates access to observing surgeries.

These challenges underscore the need for improved technologies for real-time 3D viewing during surgical procedures, particularly in resource-limited settings.

1.2 Research question

The research question guiding this study is: *"How can stereoscopic imaging techniques be optimised for real-time remote supervision during surgical procedures, and what is their impact on guidance, education and expert assistance in different applications?"*

This question includes exploration of technical optimisations for capturing and transmitting stereoscopic images, as well as assessment of their broader implications for improving surgical guidance, educational opportunities and expert assistance. The aim is threefold.

First, the aim is to develop techniques for capturing, synchronising and transmitting stereoscopic images from surgical environments to remote locations. This includes the integration of imaging systems in endoscopic devices and wearable glasses with cameras. This is done by using stereoscopic imaging with two cameras capturing images from different angles.
In second place, the focus is on implementing live streaming capabilities to enable external observers to interactively view surgical procedures in 3D, facilitating supervision, educational sessions and expert consultations. The stereoscopic video stream will be broadcast over the internet so that people can easily watch from anywhere. It must be said that patient privacy is guaranteed.
Finally, it evaluates the effectiveness of the developed systems in improving surgical supervision, education and accessibility. This evaluation includes assessing factors such as the quality of the transmitted stream.

2 Theoretical Framework

2.1 3D Viewing: Stereoscopic Vision in Humans

Human eyes are positioned roughly 50 mm to 75 mm apart, allowing each eye to perceive a slightly different angle of the whole scene. The image seen by one eye is similar to that seen by the other eye, but with a small shift. This is called parallax. To perceive depth, two images slightly different from each other are needed, as is the case with eyes. With only one eye, it is possible to only see a two-dimensional image, without depth perception. For a three-dimensional image, however, depth perception is essential. Therefore, the scene must be captured by multiple cameras from different angles and positions. A technique known as stereoscopy. [8]

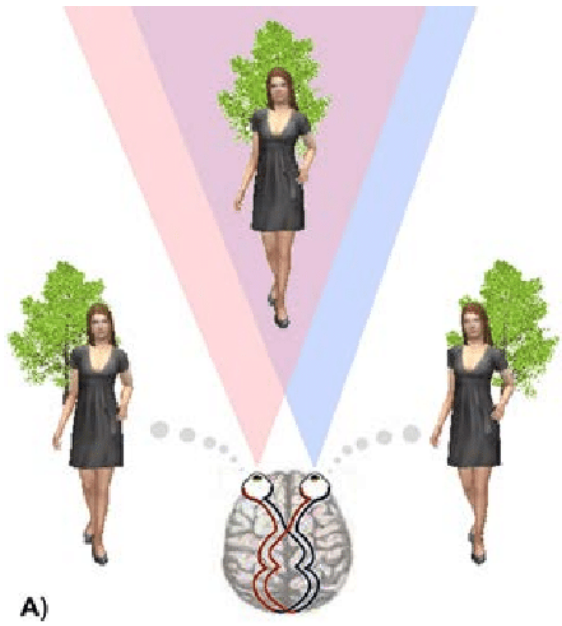


Figure 1: *Illustration of binocular vision and depth perception. The left image represents what the left eye sees, the lady to the right of the tree. The right image represents what the right eye sees, the lady to the left of the tree. The middle image shows the combined perception of depth, with the lady appearing in front of the tree. [8]*

Binocular disparity is the difference in the images seen by each eye because of their slightly different positions on the face. These differences enable the brain to perceive depth and estimate how far away objects are from us. In the primary visual cortex, also known as V1, there are specialised cells that respond to this binocular disparity. Neurons in this area can detect the differences between the images received by each eye, but they do not provide information about how deep an object is. Instead, they focus on finding the absolute disparity, which simply means how much the images differ from each other.
To perceive depth, the brain must combine this information about binocular disparity with other visual cues, such as size, shadow and perspective. This happens in other higher parts of the brain. Neurons in these areas respond to complex visual stimuli and play a role in refining depth perception. These areas are involved in forming a holistic image of the spatial environment by combining different depth cues. The interaction between different parts of the visual cortex is important in creating an accurate perception of the spatial environment. [9]

To see objects sharply at different distances, eye lenses can accommodate and verge the eyes. Accommodating is the flatter or convexity of the lens. Eye lenses can change shape to adjust focus to different distances. When looking near, the lens should become more convex and when looking far, the lens should become flatter. In vergence, the eyes work together and move in different directions to perceive depth. There are two types of vergence. Convergence is turning the eyes inwards to see objects close by. Divergence is turning the eyes outward to see an object far away.

The active angle of the eyes, the vergence angle, is the angle at which the eyes turn to focus on an object. When an object gets close, the eyes converge and the vergence angle becomes large. This is because the viewing axes of the eyes must intersect to see the nearby object sharply. When an object is further away, the eyes diverge and the vergence angle becomes small. The axes of vision are more parallel to each other because the object is at a distance. The viewing axes can be seen as dotted lines in figure 2. Furthermore, it can be seen how the vergence angle changes when an object changes distance. [10]

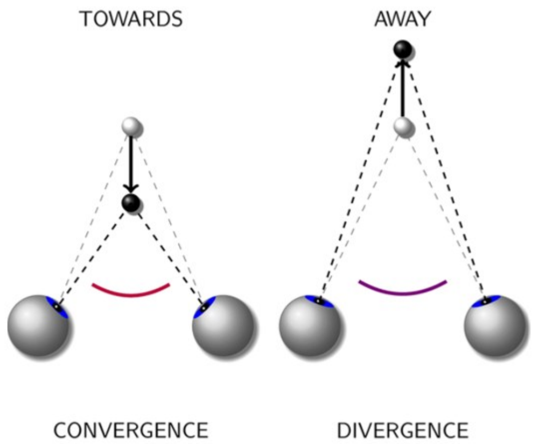


Figure 2: *Illustration of vergence angle dynamics. On the left, convergence is illustrated as the object moves closer, increasing the angle. On the right, divergence is shown as the object moves farther away, reducing the angle. [11]*

Stereopsis is the ability to perceive depth. This is due to binocular vision where each eye sees a slightly different image. Developing normal stereopsis requires the viewing axes of both eyes to be properly aligned, both eyes to have similar visual acuity and the eye muscles to be precisely controlled. Perceptual training is a method to improve 3D perception through targeted repetitive exercises. This involves performing specific visual tasks, such as recognising subtle differences in depth. Another way to improve 3D perception is to play video games, such as action games in 3D or watching films in 3D. [12]

2.2 Stereoscopic Setup

The stereoscopic system used in this project is based on a pinhole camera model and a parallel geometry. The cameras are arranged side-by-side, similar to the human eyes. The stereoscopic geometry can be seen in figure 3. This setup has the advantage of being relatively easy to build and there are no differences in light exposure and color between the two images. [13]

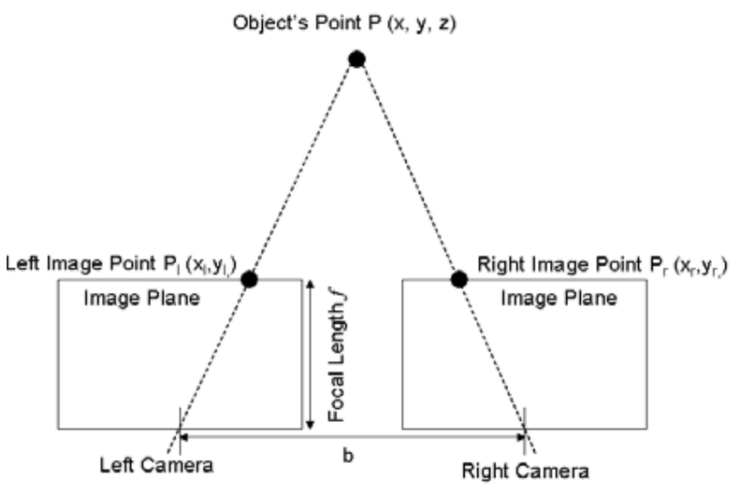


Figure 3: *The stereoscopic system used in this project is based on a parallel structure, with cameras arranged side-by-side similar to the human eyes. [14]*

2.3 Technical Requirements and Challenges

Remote real-time 3D viewing during surgery brings many advantages, but it also involves some technical challenges. The flowchart 4 shows the steps of the 3D video streaming process. This section outlines requirements to be met for each step and explains some of the technical challenges to be considered.

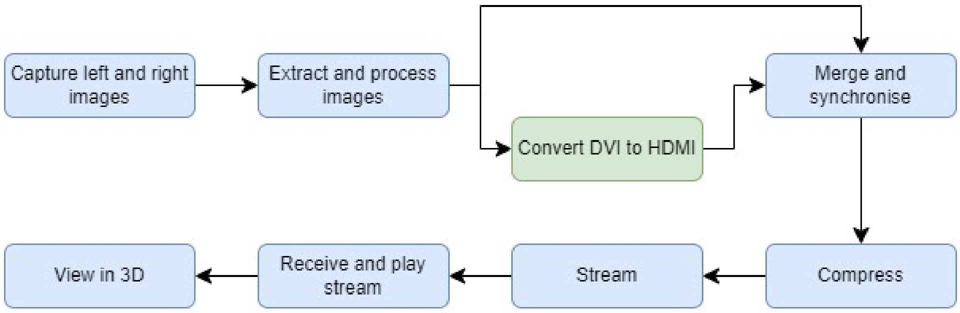


Figure 4: *Flowchart of the steps involved in 3D video streaming for remote viewing using an endoscope or camera-equipped glasses. The step in green applies only to the endoscope. For the glasses, this step is not needed and the process continues directly.*

2.3.1 Capture Left and Right Images

This part of the process requires two separate cameras that can be used for a stereoscopic setup. In the case of the endoscope, it should be a stereoscopic one with two separate cameras.
The cameras must be able to simultaneously record images from different positions. Therefore, two separate video channels are needed: one for the image of the left camera and one for the image of the right camera. This is necessary to later combine the images into a stereoscopic image. The cameras must also have high spatial resolution and be able to provide accurate colour representation.
So here are the technical challenges. When recording the left and right images, high resolution and accurate colour reproduction are essential to ensure image quality.
Another challenge is with camera-equipped glasses, where it is essential that the cameras are securely mounted and not easily displaced on impact. However, they also need to remain adjustable to ensure precise alignment, which is critical.

2.3.2 Extract and Process Images

To extract and process the images, a video processor is needed in the case of the endoscope. Camera-equipped glasses require adapters to connect the USB cameras to the laptop.
The images need to be received, processed and transmitted to the laptop. For the endoscope, the images are in DVI format, requiring a video capture device to convert the images. For the glasses, this intermediate step is not necessary and the laptop can already process the USB input.
There should be minimal to no quality loss during the processing process. Also, both images should be processed at the same time so that the left and right images remain the same to maintain stereoscopic viewing.

2.3.3 Convert DVI to HDMI

This step is only needed for the endoscope. A video capture device is needed to convert the images from DVI to HDMI to allow the laptop to receive the images. In combination with the device, software specifically designed to process images coming in through the device is required. This software, StreamCatcherPro, is an intermediate programme to bring the images in on the laptop for further processing later.
The device must be compatible with HDMI to ensure the images can be sent to the laptop. In addition, the device must support high resolutions and frame rates of the images.
The technical challenges in this step are converting the images without losing quality. There should also be no distortions during the conversion.

2.3.4 Merge and Synchronise

Merging the left and right images and placing them next to each other in a side-by-side configuration is necessary for stereoscopic viewing. If necessary, the images must be cropped so that the left camera captures what the left eye would see and the right camera captures what the right eye would see. The images must also be space properly.
The difficulty of image synchronisation lies in merging the images from both cameras into a consistent stereoscopic image. This requires precise matching and processing of the images in the software. To achieve this, the images from both cameras have to be placed side-by-side in the OBS Studio program.
To edit the images, a laptop with sufficient computing and graphics processing power is required. The programme OBS Studio should be installed on this laptop.

2.3.5 Compress

This step requires compressing the video stream to reduce the required bandwidth. The compression should be as efficient as possible to minimise quality loss. A laptop with sufficient computing power and OBS Studio installed on it is needed to execute this step. This software uses codecs to compress the video in real-time to a smaller file size without noticeable loss of quality. In addition, reliable network connections are needed, preferably 4G/5G networks or wired LAN connections that provide sufficient bandwidth.

A technical challenge in streaming high-resolution videos is the limited ability of networks to handle large amounts of data simultaneously. Streaming requires significant bandwidth to transmit and display videos smoothly and without interruptions. Network bandwidth specifies the maximum capacity of a communication link to transmit data over a network connection within a certain time period. This is represented in number of bits per second that can be transmitted.
The higher the bandwidth, the more data can be transported and received at the same time. However, bandwidth is not unlimited. This is due to several reasons. Sometimes there is limited router capacity, but it may also be because multiple devices have to share the bandwidth, which leads to a decrease in performance. Although 3G networks have significantly lower bandwidth compared to 4G networks, research has shown that 3G is sufficient for streaming live videos. Given that the average delay on 3G is 119.3 ms compared to 54.4 ms on 4G [15]. In hospital settings, however, higher bandwidth networks like 4G or 5G are preferred to ensure reliable streaming capabilities.
Compressing data is a solution to reduce the required bandwidth. By removing less important information from the data, the file size can be decreased without sacrificing its quality. [16]

2.3.6 Stream

In streaming, it is important to send the video stream from the laptop to a smartphone. This is done through the TeamLink programme installed on the laptop where OBS Studio is also installed. The stream should be smooth and without interruptions and there should be minimal loss of quality. This requires an internet connection with sufficient bandwidth. Low latency is essential for optimal cooperation between medical staff.

One of the biggest challenges in video streaming is latency. Latency is the delay between performing an action and perceiving the result of that action. This time delay can occur due to various factors, such as signal processing, data transfer between devices, image acquisition time and display technology.
In interactive applications, such as remote viewing of surgical procedures, even a small delay can seriously affect the user experience. In hospital settings, a latency of less than half a second is considered acceptable [17]. Excessive latency between instructors and remote participants can lead to communication problems, confusion or frustration. Achieving low latency is important when transmitting stereoscopic videos, especially for streaming for telementoring. Telementoring involves remote guidance and instructions between medical professionals [18].
However, live streaming technology causes some delay. This streaming delay can vary depending on many factors such as internet speed. This can be one-way, only from sender to receiver, or round-trip, from sender to receiver and back. To improve communication, the latency should be kept as close to real-time as possible. [19]

2.3.7 Receive and Play Stream

The video stream should be received and played on the smartphone in preparation for display in a VR headset. The video should be played with minimal latency and no loss of quality. There are some requirements for

the smartphone. The TeamLink app must be installed and the smartphone must be small enough to fit in the VR headset. Phones between 11.9 cm (4.7 inches) and 17 cm (6.7 inches) are suitable for the Renkforce RF-VRG-200, used in this project [20]. Also, the smartphone must have a screen with sufficient resolution to display details clearly [21].
The technical challenges associated are usability, power supply and battery life. The system must be easy and intuitive to use. Another challenge in remote real-time 3D viewing is power supply. Wireless systems are preferred by surgeons because of the added mobility and flexibility. It allows surgeons to move freely without being restricted by cables or other physical connections. However, not all devices are wireless, such as the camera-equipped glasses used in this project. If the system does require wires, a possible solution is to make them as thin and flexible as possible. Another option could be to strategically place the wires so that they do not hinder operations.
Wireless systems do rely on batteries. It is important to ensure that battery life is sufficient for prolonged operations. This requires efficient power management and the availability of spare batteries.

2.3.8 View in 3D

To see the smartphone's side-by-side images in 3D, a VR headset is needed. These VR glasses should be comfortable enough to wear for several hours without dropping off or hurting. It should have the ability to adjust the pupil and lens distance according to individual preferences. The VR headset should fit a phone, such as the Renkforce RF-VRG-200.
In addition, the systems must be user-friendly for both surgeons and remote viewers. Surgeons have busy schedules and often experience high workloads [22], so there is not much time to start up and use a difficult system during a surgical procedure. This also reduces the risk of distractions during surgery. The software should be intuitive and reliable, allowing the surgeon to concentrate on the surgery rather than technical difficulties.
For remote viewers, such as other medical professionals or students viewing the images, it is also important that the system is easy to use. Quick access to the images and easy navigation through the software are essential to facilitate efficient decision-making and assistance. It is also helpful if little to no training is needed to use the system, so that implementation of remote real-time viewing systems is as quick and easy as possible.

2.4 Ethical and Legal Challenges

Remote real-time 3D viewing in hospitals faces several legal and ethical challenges. It allows healthcare providers to watch medical images of patients remotely, allowing specialists to collaborate better and potentially improving the quality of care. However, when using this technology, healthcare institutions have to take into account various laws and regulations. In addition, there are ethical considerations related to sharing medical information, such as protecting patients' privacy.

The Dutch law 'additional provisions on processing personal data in healthcare (aanvullende bepalingen verwerking persoonsgegevens in de zorg)' in chapter 3a 'electronic processing of data', article 15a describes that the healthcare provider may only share data of a client through an electronic system, such as an electronic patient record (EPR), after it has been verified and established that the client has specifically given consent for it. This means that the client must explicitly agree to the sharing of his/her data. A client has the right to withdraw his/her consent at any time. Additionally, the healthcare provider is only allowed to share data of a client through an electronic system if the privacy of other involved parties is not affected. [23]

Since 25 May 2018, the General Data Protection Regulation (GDPR) has been in force for the entire European Union. In the Netherlands, this regulation is known as the Algemene Verordening Gegevensbescherming (AVG). The GDPR replaces the Personal Data Protection Act and describes the rules that governments, businesses and associations must follow when processing personal data. The purpose of the GDPR is to protect individuals when processing their data and to ensure the free flow of personal data within the European Economic Area (EEA). This includes all EU countries including Liechtenstein, Norway and Iceland [24].
In article 32 of the GDPR, it is described that healthcare institutions must take precautions to protect personal data. This includes implementing strong security protocols, encrypting data during transmission and storage and setting up access controls to ensure that only authorised professionals have access to medical and personal information. [25]

According to the GDPR, the processing of personal data must be done according to certain guidelines [26]:

1. Lawfulness, fairness and transparency
2. Purpose limitation
3. Data minimisation
4. Accuracy
5. Storage limitation
6. Integrity and confidentiality

These principles are explained below.

1. Lawfulness, fairness and transparency: Personal data should only be processed if there is a valid reason for doing so. It must be clear why the data is being processed and how this is being done. According to the GDPR, you may only process personal data for specific, explicitly defined and justified purposes. Furthermore, any processing of personal data must be proportionate and may only take place if there is no other way, where less or no data is collected, to achieve the purpose. For remote 3D viewing, patient data should be anonymised wherever possible to protect their privacy. The patient must give consent, explaining how their data will be used and for what purpose.

2. Purpose limitation: Data processing must serve a specific purpose. No personal data may be processed if no purpose is defined. For example, it is not allowed to collect data in case it might be useful in the future. Personal data may, however, be collected for several purposes at the same time, as long as these purposes are clearly defined in advance. The purpose must be justified, where it is important that collecting the data is necessary for the purpose being pursued. In the context of remote 3D viewing, personal data must be processed specifically to improve surgical outcomes through expert remote guidance. Additionally, these data may also serve educational purposes.

3. Data minimisation: Personal data may only be processed to the extent strictly necessary for its intended purpose. Data must be adequate, relevant and limited to what is essential. This means collecting only those data that are actually needed. The collected data must be relevant to the specific purpose and no more data than strictly necessary should be gathered. This principle requires that only essential data, such as real-time surgical images, be collected and processed.

4. Accuracy: All personal data being processed must be correct and up-to-date. This means that organisations must ensure that data is accurate at the time of collection and throughout the processing period. If data is no longer correct or up-to-date, for example because it is no longer accurate or is outdated, appropriate action must be taken. This can be achieved by deleting or correcting the data.

5. Storage limitation: The GDPR states that personal data must not be stored for longer than necessary for the purpose for which it was collected. The person whose data is collected must be aware of the retention period. Medical record data is stored for a minimum of 20 years. There may always be reasons for a longer retention period. [27]

6. Integrity and confidentiality: Data must be properly secured and remain confidential. The European and international standards NEN-ISO/IEC 27001, NEN-ISO/IEC 27002 and NEN-EN-ISO 27799 explain how healthcare providers should secure personal data. NEN 7510 is the Dutch version of this standard, which is based on the international standards. Hospitals can reduce the chances of a data breach by encoding personal data. This is very important these days, as large companies often face cyber attacks. Risk management plays an increasing role in IT security. Data encryption is a solution for hospitals. [28]
Encryption involves encoding data that is written in text form. This code will only become readable again if you use the right key. Encryption reduces the risk of a hack because the encrypted content cannot be read by third parties. It is also more secure within the company because only authorised people can access it. [29]

The controller is responsible for ensuring compliance with data protection principles. This means that the controller must be able to demonstrate that they comply with the requirements of the GDPR at all times.

The controller has the responsibility to determine how and why personal data is processed. This includes establishing the purposes of the processing and determining the methods used to achieve those purposes. [25]

The Medical Device Regulation (MDR) has existed since 26 May 2021 and is the legislation that regulates medical devices in the European Union. If 3D remote viewing is used for remote advice and guidance, as in this case, it is considered a medical device. If so, the system must conform to the MDR guidelines. The medice device must be classified depending on the potential risks of the device. The classes range from I (non-invasive and low risk), IIa (temporarily invasive and/or moderate risk), IIb (long-term invasive or high risk) to III (highest risk).

Since the 3D remote viewing system from this project is used to make decisions, it falls into a higher risk class. This is because of its direct impact on surgical procedures, patient health and patient safety. The system falls into class IIb because it poses a significant risk given the potential impact on the patient if something goes wrong with the display of the images.
The system must be assessed by a notified body for compliance with the requirements of the MDR. This requires extensive technical documentation and a clinical evaluation demonstrating the safety and effectiveness of the system. It also requires a post-market surveillance plan to continuously monitor the system after it is released on the market.
[30]

2.5 Clinical Applications

Using three-dimensional stereo imaging technology, it is possible to obtain a more realistic view of depth than traditional two-dimensional imaging technology. This improves the visualisation of anatomical structures in the human body. In the medical world, various medical imaging techniques, such as stereo endoscopy and stereo microscopy, are used to improve the precision of surgical procedures and patient safety. [13]

3D viewing technology is particularly valuable in microsurgery, where precise visualisation is needed. For example in ENT surgery, neurosurgery, stereoendoscopy and plastic surgery, 3D imaging improves the surgeon's ability to navigate complex anatomy.

During an endoscopic procedure, the doctor can look into the internal organs to diagnose and/or treat certain conditions. An endoscopy is usually done under sedation or general anaesthetic. A thin, flexible tube, with a light and a camera at the end, is inserted into the body through a natural opening. This can be seen in figure 5. The camera can take images of the internal organs that can be projected on a monitor. [31]

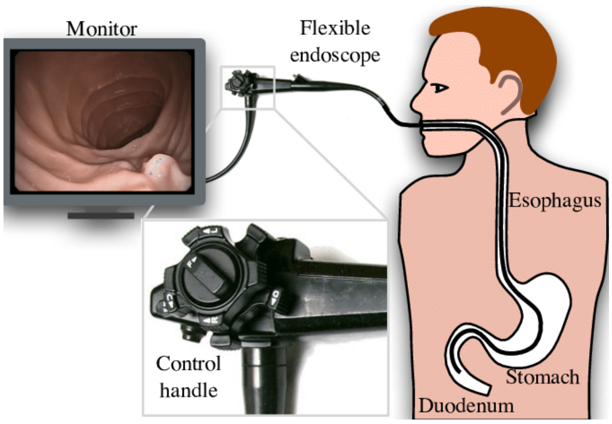


Figure 5: *Insertion of a flexible endoscope through the mouth for endoscopy. The control handle can be used to navigate through the patient's internal organs, while viewing endoscopic images on the monitor. [32]*

Surgical microscopes (figure 6), or operating microscope, have become indispensable in modern operating rooms because of their adjustable magnification and bright illumination. They provide clear visualisation of the surgical field and are especially useful in microsurgery, such as ear, nose, and throat (ENT) surgery [33], where surgeons need to be able to accurately observe small anatomical structures. [34]

Figure 6: *A surgeon looks through the eyepiece of an surgical microscope at the surgical area, which is also visible on the monitor. [35]*

Camera-equipped glasses (figure 7) are glasses with a camera attached to each temple. A surgeon can wear the glasses, allowing an external observer, who is not present in the operating room, to see exactly what the surgeon is doing. This can be used for expert assistance or for educational purposes.

Figure 7: *A person wears camera-equipped glasses connected to a laptop via wires. The cameras are attached to the temples of the glasses with clips, and additional stability is provided with insulation tape.*

Since the 1970s, several 3D imaging technologies have been developed, like computed tomography (CT) and magnetic resonance imaging (MRI). These techniques can create detailed images of parts of the body by scanning them in thin slices and then displaying the 3D structure of the body using advanced image processing algorithms.
In addition, stereoscopic imaging technologies are increasingly used in various applications inside and outside medical diagnostics. For example, it is used in education to give students a better understanding of human anatomy. It is also applied in digital mammography to detect breast cancer, in diabetic retinopathy screening to examine blood vessels in the retina of the eye and in minimally invasive surgery (MIS) to give surgeons a better view during operations. [13]

In MIS procedures, a video camera is inserted into the body through small incisions, allowing the surgeon to view the surgical area on a monitor. This technique allows surgeons to perform surgeries without having to directly look through large openings, resulting in less tissue damage, faster recovery and less pain

for the patient.
The use of a stereoscopic setup in MIS procedures makes it possible to see depth in the body. This improves the surgeon's visualisation and precision. [36]

In virtual reality (VR), two slightly different images are projected to the user's two eyes (figure 8), allowing them to perceive a three-dimensional image due to binocular disparity. These side-by-side images are sent from a laptop to a smartphone place in a VR headset, which acts as a stereoscope. The live images can be streamed to viewers worldwide, who can simply use their smartphones and VR headsets to see the operating surgeon's view. [37]

Figure 8: *Image showing a stereoscopic view in a VR headset, where two slightly different images are projected to each eye to create a 3D image. [38]*

3 Technical Aspects

3.1 Hardware and Software

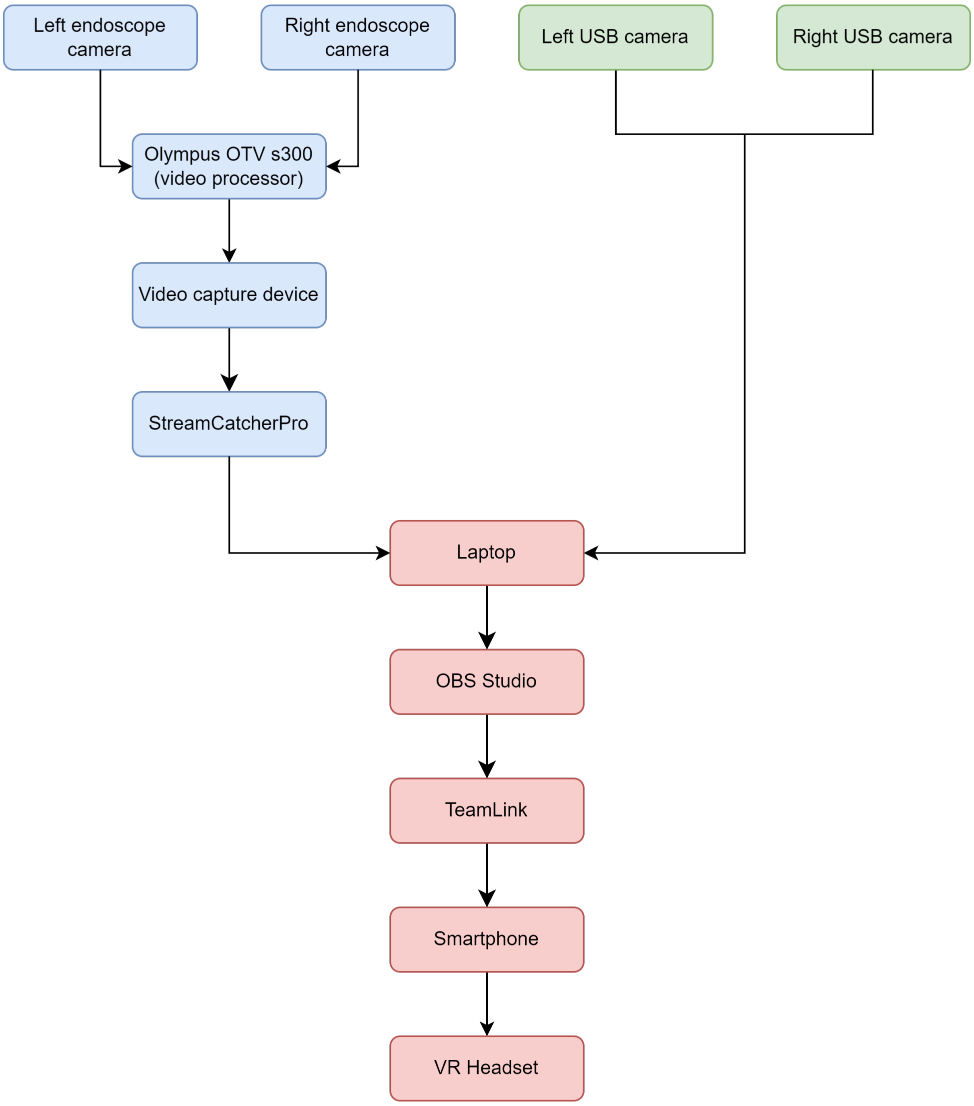


Figure 9: *Flowchart showing the hardware and software steps from camera to smartphone in a VR headset, with separate paths for endoscopy (left) and camera-equipped glasses (right).*

Figure 9 shows the flowchart of the system. The blue blocks are components needed only for the endoscope, while the green block represents the camera-equipped glasses. The red blocks are applicable to both systems. The blocks indicate both hardware and software elements.

3.1.1 Cameras

There are two types of cameras used in this project: the stereoscopic endoscope cameras and the USB cameras for the camera-equipped glasses. For the stereoscopic endoscope, the Olympus LTF-s300-10-3D (Olympus Endoeye Flex 3D) was used, which can record 3D images in HD. This videoscope has a flexible tip to view anatomical structures properly. The USB cameras used, the Groudchat JP1DV1, can record in Full HD quality. These are cameras can be connected with a wire to the USB ports of the laptop. These cameras are mounted on the glasses.

3.1.2 Endoscopic System

The blocks video processor, video capture device and StreamCatcherPro are part of the endoscopic system. To obtain side-by-side images from the endoscope, the 3D image format must be set to '3D SIDE BY SIDE'. The endoscope's cameras (Olympus LTF-s300-10-3D) record images that are processed by the Olympus OTV s300 videoprocessor. This is a device that prepares video images for display on a monitor or laptop. The next step is to send the images to the laptop. To convert the DVI connection to HDMI connection, a video capture device is used, such as the Startech USB3HDCap.

3.1.3 Laptop

To process the images, both StreamCatcherPro, Startech's programme, and OBS Studio are installed on the laptop. StreamCatcherPro manages the video capture, while OBS Studio is used to place and crop the left and right images side-by-side. The virtual camera is initiated via the StreamCatcherPro program. These images are then opened with OBS Studio, which also starts the virtual camera. The cameras of the camera-equipped glasses (Groudchat JP1DV1) also record images. This chain consists of fewer steps. The images are brought in on the laptop and edited by using OBS Studio. Streaming software TeamLink is launched on the laptop and OBS Studio's virtual camera is chosen as the camera. It is important to switch on the 'stream in HD' function to ensure the required quality.

The placement of the images in OBS Studio, such as the distance of the images from the centre, depends on the distance between the cameras. The goal is to see the same image through the VR glasses as through the endoscope or through the glasses themselves. This was achieved through repeated tests and adjustments, like trial and error. The results are shown in figure 11.
The cameras of the endoscope are located right next to each other, so the images in OBS Studio are also place next to each other. However, the cameras of the glasses are much further apart than the distance between human eyes. This creates an area that the cameras cannot see, because the cameras' images do not overlap, up to about 30 cm away from the glasses. To obtain a good image, the images are shifted closer to each other to reduce the difference between the two images.
The area between your eyes, or in this case the cameras, where you see nothing, is also shown in figure 10. This is a zoomed in version of figure 1. The vision of both eyes is shown and in between is an area, shown inside the green circle, where you cannot see any depth. Because the cameras are further apart than the eyes, this area is a lot larger. This area is called the 'human natural blind spot'. You don't notice this gap because the brain combines information from both eyes and fills in the gaps. This blind spot has been tried to be as small as possible by placing the cameras at such an angle that they are directed more towards the centre. [39]

3.1.4 Smartphone

Streaming software TeamLink sends the HD-quality images from the virtual camera of OBS Studio to the phone inserted in the VR headset (Renkforce RF-VRG-200) for a stereoscopic effect. The TeamLink app is launched on the smartphone and a connection is made to the laptop used to stream and view the footage through the VR headset.

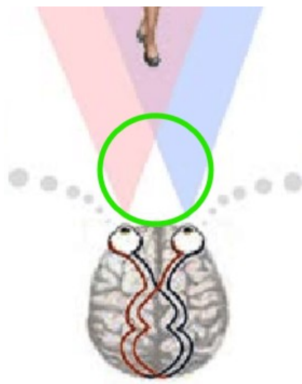


Figure 10: *Representation of the natural blind spot in the human visual field. This area is visible as a white triangle within the outlined green circle. [8]*

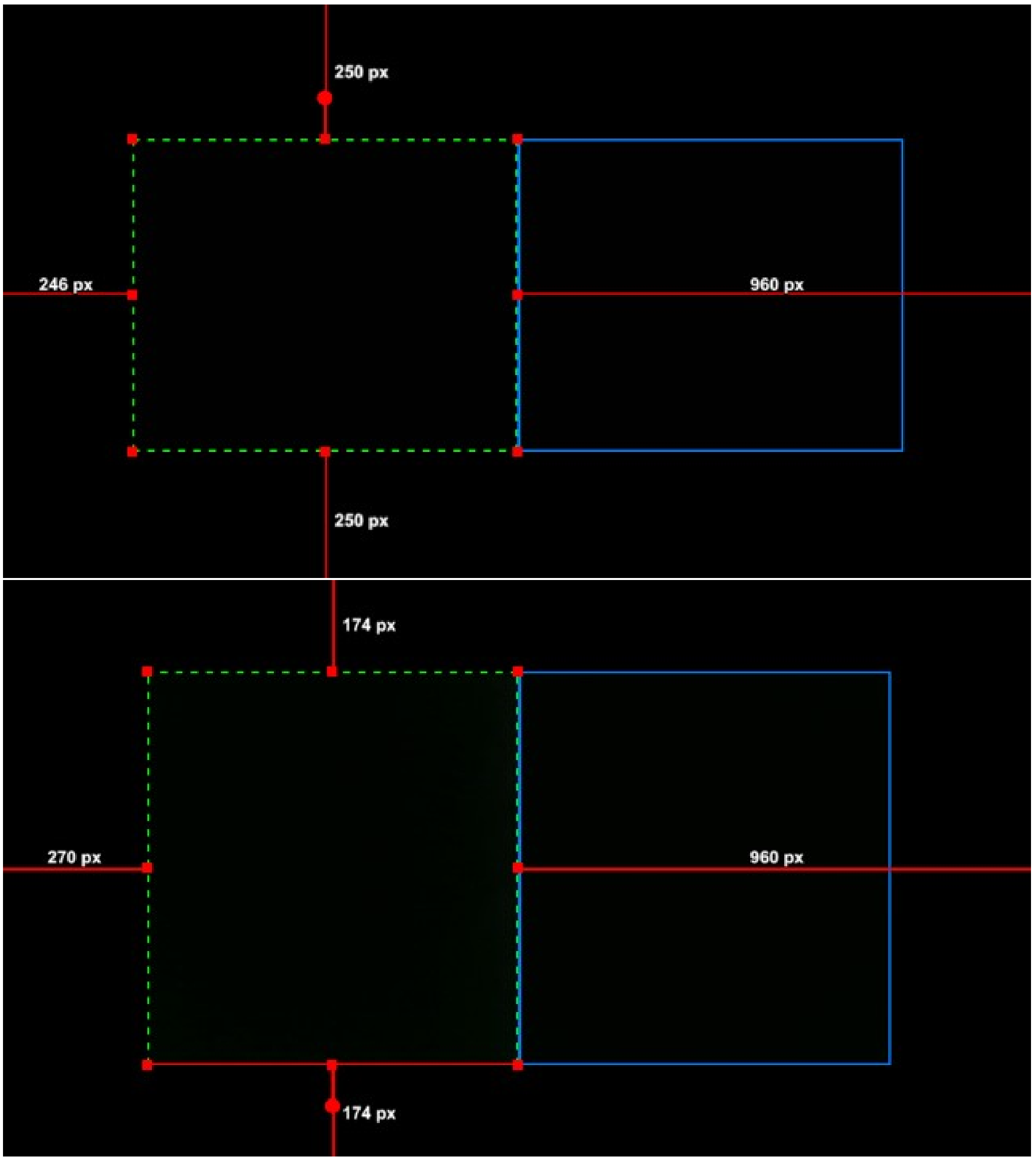


Figure 11: *Screenshots of OBS Studio settings for the cameras of both the endoscope (above) and glasses (under). The image of both settings shows an overview of the side-by-side placement of the images.*

3.2 Test Criteria for 3D Visualisation

Effective remote 3D visualisation relies on several key technical aspects that are essential for achieving optimal results. These aspects encompass various criteria that directly influence the quality and usability of stereoscopic images. Understanding and addressing these technical considerations are crucial for ensuring a seamless and immersive viewing experience in applications such as surgical supervision and remote collaboration. This section introduces the aspects tested in this project.

The test criteria and associated program requirements against which the systems are evaluated are listed in a table below.

Table 1: *Overview of the test criteria and program requirements for effective 3D visualisation.*

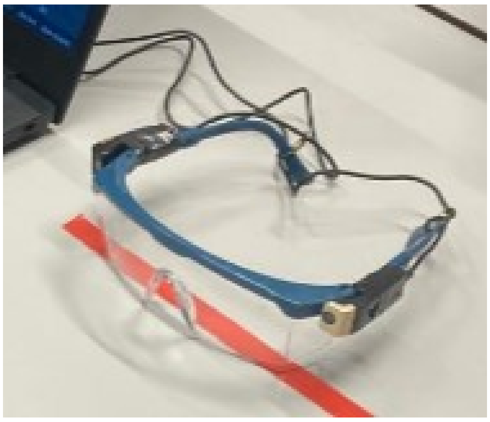
Criteria	**Program Requirements**
Resolution	The camera-equipped glasses should offer a minimum resolution of 1 lp/mm at a distance of 50 cm. The endoscope should also offer a minimum resolution of 1 lp/mm but at a distance of 7.5 cm.
Colour Accuracy	The average colour differences (ΔE) between measured and actual RGB values should not exceed 25 for both systems.
Depth of Field (DOF)	The endoscope should have a depth range of 0 cm to 10 cm measured from the camera.
Image Distortion	Distortion of the image should not be perceived as interfering during viewing.
Latency	The average latency for both systems should not exceed 0.5 seconds.
Depth Perception	In the tests for depth perception, participants should not deviate from results without the use of the VR glasses by more than 50%.

4 Methods

4.1 Technical Design

Technical design is the basis of effective 3D remote visualisation. The section 3.1 describes the hardware and software components required for the entire chain of 3D visualisation. These components are required for capturing, processing and transmitting stereoscopic images in real-time.

When designing camera-equipped glasses, several prototypes were made as an attempt to improve the glasses. Different options for mounting the cameras to the glasses were experimented with. Various laboratory glasses were considered and tried. One of the first prototypes is shown in figure 12a. In figure 7, this prototype is worn. The cameras were attached to the temples using the clips that were already attached to the cameras and insulating tape. Besides insulating tape, adhesive tape and clips were also considered and tried. The final design used cable ties and small nails to hold the cameras in place. The final design can be seen in figure 12b.



(a) *Early prototype of camera-equipped glasses. The cameras are attached to the temples using the camera clips and insulating tape over them.*

(b) *Final prototype of camera-equipped glasses, using cable ties and small nails for optimal camera positioning.*

Figure 12: *Pictures of early and final prototypes of camera-equipped glasses.*

The design and configuration of the endoscope required no modification because a pre-existing system was used.

4.2 Validation

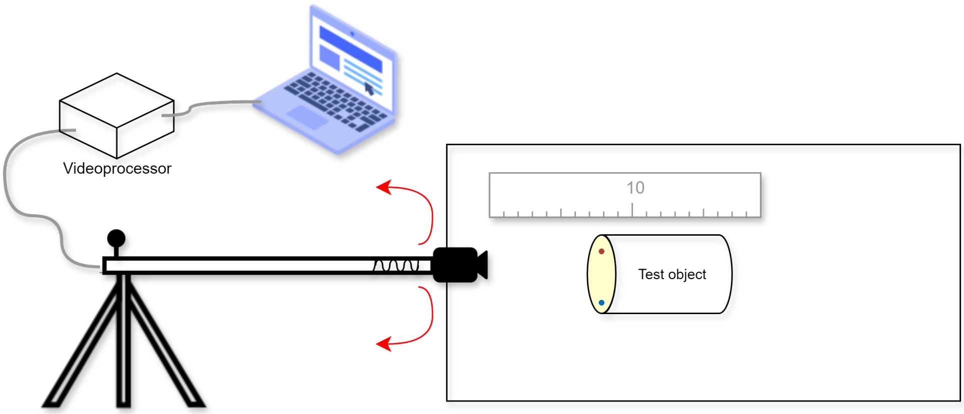
This section describes several tests performed to evaluate the quality and effectiveness of stereoscopic images taken with both camera-equipped glasses and an endoscope. The tests were designed to examine various key aspects of remote 3D visualisation, as mentioned in section 3.2. Various phantoms were used in these tests to simulate real-world conditions and ensure comprehensive evaluation across different scenarios.
The purpose of these tests is to gain deeper insights into the performance of stereoscopic images. By evaluating the images across various criteria, the aim is to obtain quantitative results. These results can then inform improvements in the technology, contributing to its further optimisation and development.
This evaluation focuses on both image quality and user experience to ensure the developed systems meet the necessary standards for effective remote 3D visualisation.

4.2.1 Test Setups

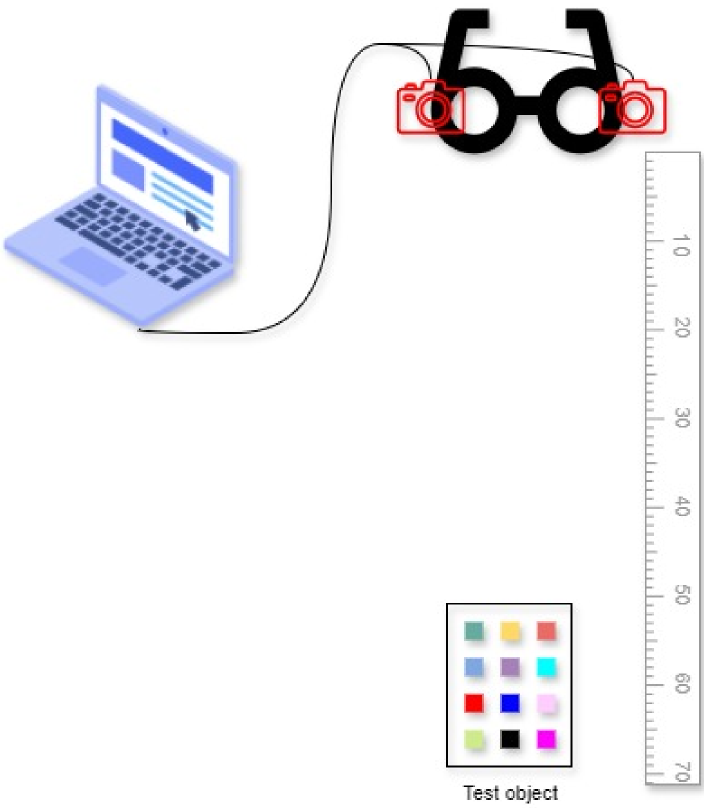
To validate the functionality and performance of both the endoscope and camera-equipped glasses, test setups were created. These setups are the basis for all tests performed.

To set up the endoscope, a tripod was used to keep the endoscope stable in the right place. The setup is recreated in a graphical representation and can be seen in figure 13a. The distances between the endoscope's camera and the test objects range from 5 cm to 10 cm. The red arrows show that the tip of the endoscope can move by using a joystick, shown as the little ball on the left side of the figure. The endoscope is connected via a cable to the video processor, which in turn is connected via a cable to the laptop.

The setup of the camera-equipped glasses is simulated and shown in figure 13b. The distances of the test objects range from 30 cm to 60 cm. The cameras on the temples of the glasses are shown in red and these are connected to the laptop via USB ports.



(a) *Graphic representation of the endoscope setup. The red arrows indicate that the endoscope head is movable using a joystick, depicted as the small sphere on the left side of the figure.*



(b) *Graphic representation of the camera-equipped glasses setup. The cameras on the glasses' temples are marked in red.*

Figure 13: *Graphic representations of the test setups for the endoscope and camera-equipped glasses.*

4.2.2 Resolution

Resolution indicates how many pixels an image or display can show. A pixel, also known as a picture element, is a point, often a square, in an image with a distinct colour. All pixels together compose the image. Resolution determines the sharpness of an image and is expressed in number of pixels that can be displayed horizontally

and vertically. The more pixels there are, the more details can be seen.
Screen size also affects image quality. An image on a smaller screen appears sharper than the same image on a larger monitor with the same resolution. This is because on a smaller screen, the pixels are closer together, making the image sharper. So larger screens need a higher resolution to maintain the same image quality as smaller screens. [40]

Spatial resolution is the ability to distinguish two closely adjacent entities as separate. It is the smallest measurable unit you can differentiate in an image. It quantifies resolution by indicating how many line pairs can be displayed per unit length. This is measured in line pairs per millimetre (lp/mm). A line pair consists of a black and a white line next to each other. The more line pairs per millimetre a system can distinguish, the better the spatial resolution and the more details can be observed. [41]

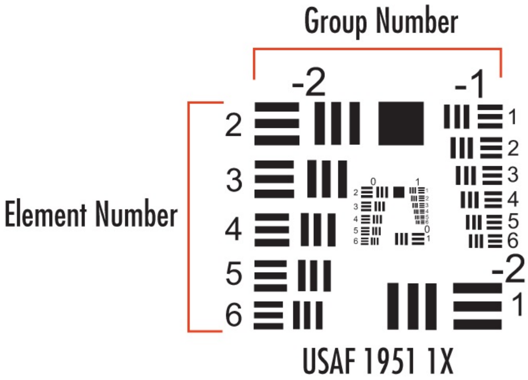


Figure 14: *USAF 1951 resolution test chart with marked groups and elements. [42]*

The USAF 1951 resolution test chart can be used to determine the resolution of a camera system. This test chart, shown in figure 14, is divided into groups. Each group has lines that are closer and closer together, divided into six elements. You look at the lines and try to see which elements are still clearly distinguishable as separate lines. The limit resolution is the group and element where the lines can still just be seen as separate elements. The formula to calculate the resolution is: [43]

$$\text{resolution} = 2^{\left(\text{group}+\frac{\text{element}-1}{6}\right)}\ \text{lp/mm} \quad (1)$$

The aim of this test is to evaluate the resolution of the image displayed through the smartphone in the VR headset to determine if it is acceptable.
The exact list of materials needed for all tests can be found in appendix C.
To evaluate the resolution of the image displayed through the smartphone in the VR headset, the test target is placed on a flat surface. The camera-equipped glasses are positioned at varying distances of 30, 40, 50, and 60 cm from the test target to reflect the range of distances a surgeon typically operates from the table. Subsequently, a setup is built where the endoscope is mounted on a tripod, at distances of 5, 7.5, and 10 cm away from the test target. This configuration allows for detailed close-up images to be captured of the test chart. By streaming the videos from both the glasses and the endoscope to the smartphone via TeamLink, the images could be viewed through the VR headset.

The resolution analysis looked at the finest patterns on the test chart that were still clearly visible. Formula 1 was then used to determine the resolution.

4.2.3 Colour Accuracy

Colour accuracy refers to how precisely the colours in an image match the original colours. The original colours are represented in a type of code, RGB values. These are the values for red, green and blue that together determine the colour of a pixel. RGB space is a colour model that describes the colour of a given point in space by the amount of red, green and blue required to produce that colour. The values for each colour range from 0 to 255, with (255, 0, 0) representing bright red, for example.

The Euclidean distance is used to measure the colour difference between two colours, namely the original coulors and the coulors in the image. The Euclidean distance is calculated with the formula:

$$\text{distance} = \sqrt{(R_2 - R_1)^2 + (G_2 - G_1)^2 + (B_2 - B_1)^2} \quad (2)$$

Where:

- R_1, G_1, B_1 are the actual RGB values
- R_2, G_2, B_2 are the measured RGB values

Using this formula, the color differences can be calculated and evaluated. [44]

The MacBeth ColorChecker (15) is a tool that can be used to check the colour reproduction of cameras or monitors. It is a card with 24 coloured squares, each having a different colour with a known RGB value. By taking an image of the map and analysing it with, for example, MATLAB, the colours can be determined and compared with the original colours.

Figure 15: *MacBeth ColorChecker used for evaluating colour accuracy. [45]*

The purpose of this test is to check whether the colours displayed on the smartphone match the original colours.
The glasses with camera are positioned at a distance of 50 cm from the test object, which is set down on a smooth underground. A setup is then built where the endoscope is mounted on a tripod, 5 cm away from the test object. Both the camera-equipped glasses and the endoscope captured images in both vertical and horizontal orientations. This was done to mitigate the potential for shadows to be cast by the upright test chart, thereby ensuring optimal image quality. By streaming the videos from both the glasses and the endoscope to the smartphone via TeamLink, screenshots could be taken of the images. The screenshots can then be analysed via MATLAB (appendix B to determine how much the colours differ from the original colours.

The number that comes out of the Euclidean distance calculation from the script indicates the colour difference between the actual RGB values and the measured RGB values. The larger this number, the greater the colour difference. A Euclidean distance that deviates less than 10.5% from the original colour is considered acceptable. This equates to 10.5% of 255, corresponding to a difference of about 27 in Euclidean distance [46]. In this case, the Euclidean distance for each of the 24 colour patches is calculated. From these values, the average is calculated. An average low value indicates good colour reproduction, while a high value indicates deviations in the system's colour reproduction.

4.2.4 Depth of Field

Depth of field (DOF) is the range of distances in a scene that are in focus. It describes the distance from the nearest to the furthest point in an image that is in acceptable focus. A proper DOF ensures that both the foreground and background of the image, in this case the surgical field, remain in focus. A good DOF also helps the eyes to refocuss less frequently, reducing eye fatigue. A better DOF also provides accurate depth perception. The purpose of this test is to determine the effective depth of field (DOF) of the endoscope when images are displayed through VR glasses. [47]

By attaching millimetre paper inside a hollow cylinder and taking video recordings from it, it is possible to assess how far objects can be away from the camera and still remain sharply visible. These measurements will help evaluate the suitability of the endoscope for accurate 3D visualisation in a surgical context. The test phantom is a hollow cylinder covered with millimetre paper. This phantom was used to perform the tests and is shown in figure 16.

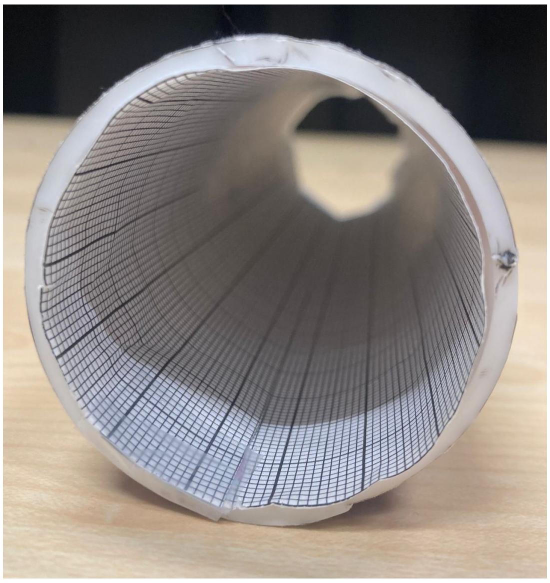


Figure 16: *The test phantom used for the depth of field measurements. This is a cylinder covered with millimetre paper to determine the maximum focusing distance. The millimetre squares serve as reference points.*

4.2.5 Image Distortion

Image distortion, also known as geometric distortion, are distortions or changes in the image caused by the optical properties of the display system. VR glasses use lenses to magnify the image, providing an immersive experience and enhanced depth perception. However, these lenses also cause distortions that can affect the accuracy and realism of the images the user sees. Distortions can cause shifts in the display of objects, making distances and dimensions appear inaccurate. This can be problematic in surgery because precision is of high importance. Determining the geometric distortion of images is the aim of this test, where straight lines may appear curved. [48]

To determine the geometric distortions in the images, a test was performed with millimetre paper. The test target was placed on a flat surface. The camera-equipped glasses were positioned at various fixed distances from the test target: 10 cm, 30 cm, and 50 cm. The endoscope was attached to the tripod at different distances from the test target, specifically 5 cm, 7.5 cm, and 10 cm. Via TeamLink, the videos were streamed from the laptop to the smartphone, allowing for viewing through the VR headset. While viewing the images, close attention was paid to any distortions in the position, shape or size of the millimetre paper.

The degree of distortion was rated on a scale of 1 to 5, with a score of 1 being assigned to images with no distortion, while a score of 5 was assigned to images with severe distortions. This rating was conducted by interviewing five individuals who examined the images and provided their assessments. The score form used for this evaluation has been included in the appendix for reference (A).

4.2.6 Latency

In section 2.3, latency is introduced as a challenge in remote real-time 3D viewing. Latency can affect the accuracy of interventions, especially when users need to respond quickly to visual input, such as during surgical procedures. The purpose of this test is to measure the delay between an action recorded by camera-equipped glasses and the endoscope, and the display of these images in VR glasses holding a smartphone.

To measure the delay between recording and displaying the images, a stopwatch was used. For the latency test, something extra was added to the setups in figure 13. The laptop was placed in view of the camera so that the camera could film the laptop's screen. An online stopwatch was started on the laptop. As the camera filmed the stopwatch image, the stopwatch could also be seen on the phone, which received the streamed image via TeamLink. For this test, the smartphone was not placed in the VR headset. An external second phone was used to create a video showing both the laptop and the smartphone, with both showing the stopwatch. The time on the actual stopwatch and the time displayed via the smartphone were then noted. By calculating the difference between these two times, the latency could be determined.

4.2.7 Depth Perception

Stereoscopic depth perception is the ability to perceive depth in a scene by processing slightly different images received by each eye. Depth perception plays a major role in creating a sense of depth and realism in virtual environments. Research has shown that the distance between the two cameras should match the user's pupil distance. This ensures that the images received by each eye coincide correctly and enable realistic depth perception.

When the cameras are closer together than the actual pupil distance, your eyes have to turn inwards more, converge more, to see the stereoscopic image. This can cause objects to appear closer than they really are. If the cameras are further apart than the pupillary distance, the eyes have to turn less inward. This lowers the amount of convergence. This makes objects appear further away than they really are. It is difficult for the brain to merge images with too much separation, which can lead to reduced viewing comfort and even double vision. [48]

The aim of the depth perception test is to investigate how effective stereoscopic images are for estimating depth in different scenarios. Two different test methods were used for this test.

For the method with camera-equipped glasses, participants participated in two tasks: playing the games 'Dr. Bibber' and a 'nerve spiral'. These games were played both with and without the VR glasses, to which the camera-equipped glasses were attached. This so that through the VR glasses you can see the game and what you are doing. The time taken to complete the games and the number of errors were measured and compared between the two conditions.
Dr. Bibber (figure 17a) and nerve spiral (figure 17b) are games that require fine motor skills and good hand-eye coordination. In Dr. Bibber, players must use tweezers to carefully remove body parts from a plastic 'patient' without touching the sides of the openings. If the player touches the edge, the buzzer goes off and thus a mistake has been made. In the nerve spiral, the aim is for players to manoeuvre a small metal loop along a metal track without the loop touching the spiral track. If it does, the beeper goes off and the player has made a mistake. So the aim of both games is to make as few mistakes as possible within a given time.

(a) *Image of Dr. Bibber, where a bone is being removed from the body with tweezers. [49]*

(b) *Image of the nerve spiral, showing the metal loop and track. [50]*

Figure 17: *Images of the games used as test objects to assess the depth perception of the camera-equipped glasses.*

In the endoscope method, a hollow cylinder was covered with paper and beads placed directly against the camera. Participants viewed the images through VR glasses and were asked to estimate the distances between the beads in the cylinder. Additionally, participants were asked to determine the order of the beads from closest to farthest. These estimations were compared with the known distances and order, allowing it to be assessed how well endoscope images can represent depth.The phantom used for this test is a paper cylinder with beads, shown in figure 16.

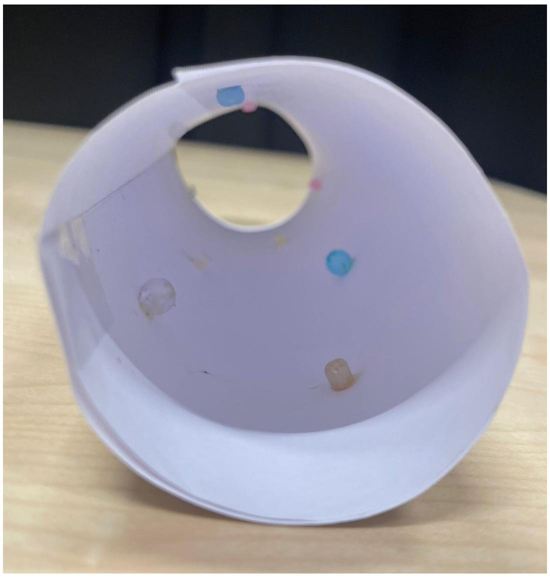


Figure 18: *Test phantom used for depth perception measurements with the endoscope. This is a paper on which beads are glued at known distances. The paper is rolled into a cylindrical shape.*

5 Results

The results of the tests provide insight on how well the remote 3D viewing system works with both the endoscope and camera-equipped glasses. In this section, collected data is analysed and the results are summarised. Each aspect is evaluated to assess the suitability of the systems. It also includes a comparison of the results found with the previously established program requirements listed in table 1.

5.1 Resolution

Table 2: *Resolution test results of the camera-equipped glasses and the endoscope at different distances.*

Distance [cm]	Group	Element	Resolution [lp/mm]
Camera-equipped glasses			
30	-1	6	0.9
40	-1	4	0.7
50	-1	1	0.5
60	-1	1	0.5
Endoscope			
5	1	4	2.8
7.5	1	2	2.2
10	0	6	1.8

Table 2 presents the resolution results obtained from tests conducted with the camera-equipped glasses and the endoscope at various distances. Resolution, expressed in line pairs per millimeter (lp/mm), serves as an indicator of the clarity and sharpness of the captured images. The resolution was calculated using formula 1.

The pre-established program requirement for resolution was a value of 1 lp/mm for the camera-equipped glasses at 50 cm and for the endoscope at 7.5 cm. As shown in the table of results above, the values found are 0.5 lp/mm and 2.2 lp/mm, respectively. This means that the glasses do not meet the requirement, while the endoscope does.

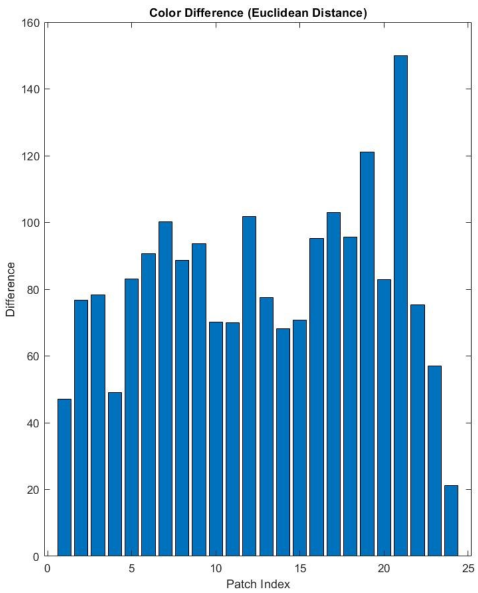
5.2 Colour Accuracy

Table 3: *Table of average colour differences and standard deviations (SD) for different test charts and systems.*

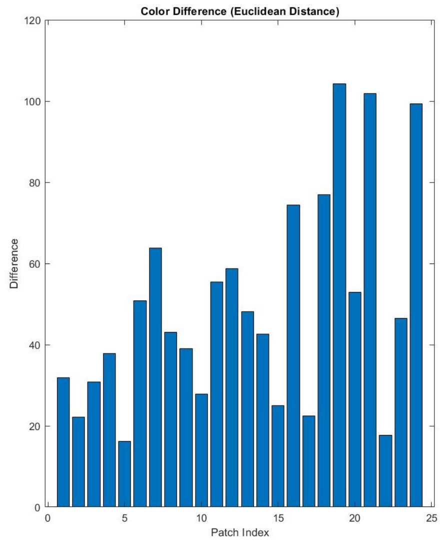
Test chart	System	Average colour difference (SD)
Vertical	Camera glasses	82 (± 25.8)
Horizontal	Camera glasses	50 (± 26.1)
Vertical	Endoscope	91 (± 34.3)
Horizontal	Endoscope	59 (± 22.8)

The colour accuracy of both systems was evaluated by comparing the measured RGB values with the actual RGB values from a MacBeth ColorChecker. This was analysed using a MATLAB script that can be found in Appendix B. The average difference, the Euclidean distance, between the measured and actual RGB values for both the standing (vertical) and lying (horizontal) test charts can be found in table 3. It was stated beforehand that the colour differences 25 should not exceed. Both systems do not meet this requirement.

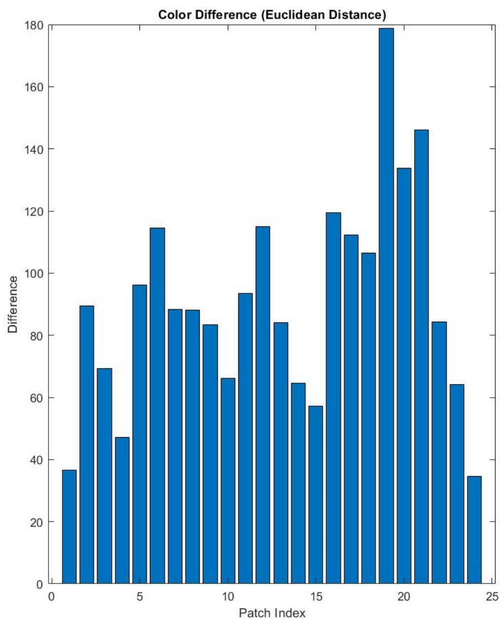
Figure 19 shows the results of measurements. The plot displays the colour differences as a bar chart, showing the Euclidean distance for each patch. This shows which patches show the largest deviations. The graphs show that some patches have significant colour differences, contributing to the relatively high mean differences. If a patch has a value below 27, the colour reproduction is still accurate.
The x-axis shows the different colour patches of the MacBetch ColorChecker. Each patch index corresponds to a specific colour patch. The y-axis shows the size of the colour difference, measured in Euclidean distance. This distance indicates how far the measured colour differs from the actual colour. The higher the bar, the greater the colour difference. The values of the individual patches were summed and divided by the total to arrive at the average values, as shown in table 3.



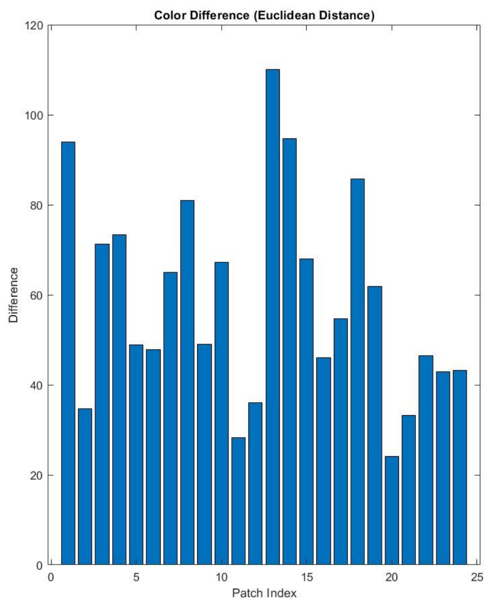
(a) *Vertical test results for glasses.*



(b) *Horizontal test results for glasses.*



(c) *Vertical test results for endoscope.*



(d) *Horizontal test results for endoscope.*

Figure 19: *Results of measurements for different systems and positions of the test chart. Comparison between actual and measured RGB values for 24 color patches (left subplot) and color differences represented as bar charts (right subplot) for different test charts and systems.*

5.3 Depth of Field

The test shows that every square up of the phantom (16) to and including the 15th millimetre square is sharply visible. Up to about the 30th millimetre square, the squares can still be distinguished, but after that the image becomes significantly blurred. Although the horizontal lines remain somewhat recognisable after the 30th square, the vertical lines fade almost completely. This implies that the effective depth range of the endoscope extends from 0 cm to 3 cm from the camera.
The stated requirement was that the depth of field should be 0 cm to 10 cm. This was partly met, because after 3 cm, squares can still be distinguished and it is not completely blurred.

5.4 Image Distortion

This test was performed by 5 participants who completed the score form (see appendix A) and also made comments.

Table 4: *Scores for the camera-equipped glasses.*

Participant	Score
1	1
2	1
3	2
4	1
5	1
Average (SD)	1.2 (±0.4)

A participant commented that the further the test target was from the glasses, the more distortion occurred.

Table 5: *Scores for the endoscope.*

Participant	Score
1	1
2	5
3	2
4	2
5	5
Average (SD)	3 (±1.7)

Three participants noted that the squares on the millimeter paper had transformed into rectangles, with horizontal stretching. However, all squares remained equal in size and shape.

Due to these disappointing scores for the endoscope, adjustments were made to the OBS settings (see figure 11). A compromise was chosen to minimise distortion while maximizing visibility. When the squares were made completely square, the image became so small that a large portion of the screen turned black.These adjustments were also demonstrated to the participants.
The new scores for the endoscope were:

Table 6: *New scores for the endoscope.*

Participant	Score
1	1
2	3
3	2
4	2
5	4
Average (SD)	2.4 (±1.0)

The requirement was that the distortion should not be interfering with the person viewing the images. An endoscope already causes distortion by itself, so in clinical practice people are used to this. A mean score for the camera-equipped glasses is 1.2, which means almost no distortion. An average score of 2.4 was given for the endoscope, which is between very slight and moderate distortion. Taking into account familiarisation in clinical practice, this requirement is thus met by both systems.

5.5 Latency

Table 7: *Latency measurements for the camera-equipped glasses and endoscope.*

System	Actual stopwatch time [s]	Displayed stopwatch time [s]	Latency (s)
Camera-equipped glasses	5.17	4.73	0.44
Camera-equipped glasses	6.16	5.78	0.38
Camera-equipped glasses	7.90	7.49	0.41
Average (SD)			0.41 (± 0.025)
Endoscope	7.44	6.46	0.98
Endoscope	9.25	8.22	1.03
Endoscope	11.40	10.33	1.07
Average (SD)			1.03 (± 0.037)

The latency of the camera-equipped glasses and the endoscope was measured by comparing the actual stopwatch time with the displayed stopwatch time. The latency was calculated as the difference between the actual time and the displayed time. The results are summarised in Table 7.

The table shows that the latency for the camera-equipped glasses range from 0.38 to 0.44 seconds, with a average latency of 0.41 seconds. While the latency for the endoscope range from 0.98 to 1.07 seconds, with a average latency of 1.03 seconds. This indicates that the endoscope has a higher latency compared to the camera-equipped glasses.

The preset requirement is that the latency should be less than 0.5 seconds for both systems. The results show that the camera-equipped glasses meet this requirement, with an average latency of 0.41 seconds. However, the endoscope does not meet this requirement, as the average latency is 1.03 seconds.

5.6 Depth Perception

This section presents the results of the depth perception tests for both camera systems. Two different methods were used for these tests, therefore the results are presented in separate sections.

5.6.1 Camera-Equipped Glasses

The tests for the camera-equipped glasses, playing Dr. Bibber and nerve spiral, were performed by 3 participants. Each subject played both games six times, three times with and three times without VR glasses. Time of playing and the amount of mistakes made were examined. Below is an overview of the results:

Table 8: *Test results of Dr. Bibber gameplay sessions without VR glasses.*

Participant	Attempt	Time [mm:ss]	Errors
1	1	02:08,45	3
	2	01:43,60	4
	3	01:00,15	1
2	1	01:51,94	7
	2	01:53,85	4
	3	02:01,44	5
3	1	01:15,19	4
	2	01:55,13	10
	3	01:12,01	4
Average Time (SD)	-	01:43,08 (±0:31,04)	4.5 (±2.8)

Table 9: *Test results of Dr. Bibber gameplay sessions with VR glasses.*

Participant	Attempt	Time [mm:ss]	Errors
1	1	05:45.93	7
	2	04:18.06	12
	3	03:16.67	7
2	1	05:08.72	13
	2	04:38.73	6
	3	06:21.01	9
3	1	04:32.17	5
	2	04:11.06	7
	3	03:41.10	6
Average Time (SD)	-	04:44.34 (± 00:45.38)	7.3 (± 2.6)

The average times and errors for the tests without and with VR glasses are converted into seconds and then used to calculate the percentage differences.

- Without VR glasses: 1 minute and 43.08 seconds $= 60 \times 1 + 43.08 = 103.08$ seconds
- With VR glasses: 4 minutes and 44.34 seconds $= 60 \times 4 + 44.34 = 284.34$ seconds
- Without VR glasses: 4.5 errors
- With VR glasses: 7.3 errors

$$\text{Percentage difference} = \left(\frac{284.34 - 103.08}{103.08}\right) \times 100\% = \left(\frac{181.26}{103.08}\right) \times 100\% = 175.85\% \tag{3}$$

$$\text{Percentage difference} = \left(\frac{7.3 - 4.5}{4.5}\right) \times 100\% = \left(\frac{2.8}{4.5}\right) \times 100\% = 62.22\% \tag{4}$$

- **Average time**: 176% increase
- **Average errors**: 62% increase

The requirement was that the deviation should be less than 50%. As the percentage differences for both the average time and the average number of errors significantly exceed this threshold, the requirement is not met.

Participants also shared their experiences. They complained of neck pain during the game because of the need to look down when playing this game. In addition, the VR glasses sagged when looking down, making participants feel they had to hold them. Depth perception also proved difficult, as one hand often blocked one of the cameras when removing the ‘bones’ from the game, resulting in loss of stereoscopic vision and making it difficult to perceive depth accurately.

Table 10: *Test results of nerve spiral gameplay sessions without VR glasses.*

Participant	Attempt	Time [mm:ss]	Errors
1	1	01:00.60	0
	2	00:51.18	0
	3	00:43.35	0
2	1	00:48.47	1
	2	00:36.50	3
	3	00:38.98	1
3	1	00:42.26	3
	2	00:41.68	3
	3	00:49.41	4
Average Time (SD)	-	00:44.49 (± 00:08.47)	1.7 (± 1.5)

Table 11: *Test results of nerve spiral gameplay sessions with VR glasses.*

Participant	Attempt	Time [mm:ss]	Errors
1	1	01:44.55	13
	2	02:32.82	7
	3	02:39.67	1
2	1	02:04.38	10
	2	02:33.18	8
	3	02:39.43	4
3	1	01:07.48	6
	2	01:18.51	10
	3	01:44.68	11
Average Time (SD)	-	01:56.29 (± 00:34.24)	7.8 (± 3.5)

The same process is followed here to convert the average times and errors into seconds and calculate the percentage differences.

- Without VR glasses: 0 minutes and 44.49 seconds = 44.49 seconds
- With VR glasses: 1 minute and 56.29 seconds = $60 \times 1 + 56.29 = 116.29$ seconds
- Without VR glasses: 1.7 errors
- With VR glasses: 7.8 errors

$$\text{Percentage difference} = \left(\frac{116.29 - 44.49}{44.49}\right) \times 100\% = \left(\frac{71.80}{44.49}\right) \times 100\% = 161.43\% \tag{5}$$

$$\text{Percentage difference} = \left(\frac{7.8 - 1.7}{1.7}\right) \times 100\% = \left(\frac{6.1}{1.7}\right) \times 100\% = 358.82\% \tag{6}$$

- **Average time**: 161% increase
- **Average errors**: 359% increase

The condition of a deviation of less than 50% has not been met.

Participants reported that using the VR glasses was tiring on the eyes as they had to constantly refocus, causing double vision. In addition, the cameras attached to the glasses shifted frequently, leading to reduced depth perception. Holding the game made participants feel better where their hands were. One participant also reported experiencing nausea from wearing the VR glasses.

5.6.2 Endoscope

The endoscope method test involved participants viewing images of beads placed within a hollow cylinder, positioned directly against the camera, through VR glasses. Participants were asked to estimate the distances between the beads and determine the order of the beads from closest to farthest. The correct answers are as follows:

1. Order of the beads: pink, yellow, blue, brown
2. Distance pink-blue: 9.5 cm
3. Distance pink-yellow: 3 cm
4. Distance blue-brown: 5.5 cm

The answers provided by the participants are detailed in table 12, along with the calculated average and standard deviation for each test.

Table 12: *Depth perception test results using endoscope method.*

Participant	Order of beads	Distance pink-blue [cm]	Distance pink-yellow [cm]	Distance blue-brown [cm]
1	pink, yellow, blue, brown	5	3	4
2	pink, yellow, blue, brown	5	3	4
3	pink, yellow, blue, brown	6	5	8
4	pink, yellow, blue, brown	7	4	5
5	pink, yellow, blue, brown	7	2	2
Correct answer	pink, yellow, blue, brown	9.5	3	5.5
Average (SD)	-	6 (± 0.8)	3.4 (± 1.0)	4.6 (± 2.0)
Average deviation	-	3.5	0.4	0.9

The percentage differences between the average estimated distances and the correct answers are calculated as follows:

- Distance pink-blue:

$$\text{Percentage difference} = \left(\frac{6 - 9.5}{9.5}\right) \times 100 = -36.84\% \tag{7}$$

- Distance pink-yellow:

$$\text{Percentage difference} = \left(\frac{3.4 - 3}{3}\right) \times 100 = 13.33\% \tag{8}$$

- Distance blue-brown:

$$\text{Percentage difference} = \left(\frac{4.6 - 5.5}{5.5}\right) \times 100 = -16.36\% \tag{9}$$

The analysis reveals that all participants correctly identified the order of the beads. However, the estimations of the distances varied. The requirement was that the deviation should be less than 50%. The percentage differences for all three distances are within this threshold: -37% for the pink-blue distance, 13% for the pink-yellow distance and -16% for the blue-brown distance. Therefore, the requirement of less than 50% deviation is met.

6 Discussion

This study started with the research question: "How can stereoscopic imaging techniques be optimised for real-time remote supervision during surgical procedures, and what is their impact on guidance, education and expert assistance in different applications?". The discussion focuses on the interpretation of the results, the clinical context, possible improvements and future perspectives.

Resolution

The resolution test results show that the camera-equipped glasses have a lower resolution than the endoscope, as shown in table 2. For both systems, the resolution decreased as the distance from the test target increased, which is consistent with expectations. This is in accordance with the properties of optical systems, where resolution decreases as the object is further away. The highest resolution at the glasses was achieved at the smallest distance of 30 cm, with a resolution of 0.9 lp/mm. The highest resolution with the endoscope was achieved at a distance of 5 cm, with a resolution of 2.8 lp/mm.

The measured resolution of the endoscope (2.8 lp/mm) is much lower than the typical resolution of 20-30 lp/mm reported for conventional ultra-high-resolution endoscopes [51]. This deviation can be attributed to several factors, such as streaming of the images. This goes through multiple intermediate devices and networks which can lead to loss of image quality. Each step in the process (9), such as compression of the images during streaming, can result in a loss of detail and sharpness in the final display of the images through the VR glasses. This may affect the accuracy of the resolution measurements performed in this study.

It was stated beforehand that the resolution for the camera-equipped glasses at 50 cm and for the endoscope at 7.5 cm should be at least 1 lp/mm. The measured values are 0.5 lp/mm for the glasses and 2.2 lp/mm for the endoscope, respectively. These results indicate that the glasses do not meet the program requirements for resolution, while the endoscope does.

Colour Accuracy

After performing the colour display test, a few things were noticed. The average colour differences for horizontal test charts are significantly lower compared to the vertical charts. This suggests that colour accuracy improves when the test card lies flat and is filmed from above. This may be due to better lighting or less shadowing in this setup. In this setup, light can be distributed more evenly across the test card, minimising shadows and improving colour accuracy. This difference is due to lighting conditions and not technological limitations. It is thus rather due to a measurement error.

It is important to mention that the accuracy of this test may be affected by the colour reproduction capabilities of the printer used to print the MacBeth ColorChecker chart. It is not known whether the printer is capable of reproducing the exact colours as specified. This may therefore lead to discrepancies between actual and measured colours.

The colour accuracy of both systems was assessed by comparing the measured RGB values with the actual RGB values from the MacBeth ColorChecker. The results are shown in table 3. According to the predefined programme requirements, colour differences should not exceed 25. However, neither system meets this requirement.

Depth of Field

The depth of field (DOF) test results show that the effective depth of the endoscope image has a range of 0 cm to 3 cm. This means that objects remain sharply visible up to 3 cm from the camera, but after that details fade rapidly. These findings are important for the practical application of the endoscope in surgical environments, where a clear and sharp view of the surgical area is crucial. These results imply that the endoscope is especially suitable for close-up images within a limited distance. This is especially inconvenient for remote viewing of surgical operations, where a wide range of distances must remain in sharp focus for optimal observation.

Manufactures claim that flexible endoscopes generally have a DOF of 1.5 mm to 100 mm [52]. This was therefore the stated program requirement, which was partially met. The DOF of 0 cm to 3 cm measured in this test is significantly lower than the upper limit of these claimed specifications. This difference may be due to several factors, such as the way the images are displayed through the VR glasses or the specific settings and conditions under which the test was performed.

The depth of field test was not performed for the camera-equipped glasses because of their different design and scope. Camera-equipped glasses are designed for wider environmental images and less for close-up details like the endoscope.

Image Distortion
The results of the image distortion test show significant differences in the degree of distortion between the camera-equipped glasses and the endoscope. For the camera-equipped glasses, the scores of the participants indicate that there was hardly any distortion, with a mean score of 1.2. This indicates that the camera-equipped glasses caused little to no geometric distortion, even at different distances from the test object.
In contrast, the endoscope initially showed much more distortion. Initial test results gave a mean score of 3 with a high standard deviation of 1.7, indicating significant variation in perceived distortion. Three participants noted that the squares on the millimetre paper turned into rectangles, with horizontal stretching observed.
To minimise this distortion, adjustments were made to the settings in OBS Studio. This resulted in a compromise where the distortion was reduced, but the image still remained large enough to be useful without blackening much of the screen. After these adjustments, the test was run again, resulting in an improved mean score of 2.4 with a standard deviation of 1.0. This indicates that the distortion had been reduced but was still present, to a lesser extent.
The program requirements were that the distortion should not be annoying to the viewer. Since endoscopes always cause a certain amount of distortion, doctors are used to this distortion . So both systems meet this requirement.

Latency
The results of the latency test show that there is a clear difference in the latency between the camera-equipped glasses and the endoscope. The camera-equipped glasses show a latency ranging from 0.38 to 0.44 seconds, with an average latency of 0.41 seconds. In contrast, the latency of the endoscope ranges from 0.98 to 1.07 seconds, with an average latency of 1.03 seconds.
The measured latency of 0.41 seconds for the camera-equipped glasses is acceptable for many applications, but may still have some impact on the accuracy and timing of interventions during surgical procedures. The endoscope, with an average latency of 1.03 seconds, shows a significantly higher delay, which can be problematic in situations where fast responses to visual input are crucial. The requirement was that the latency should not exceed 0.5 seconds. The glasses meet this requirement, but the endoscope does not.
During the test, it was also noted that the stopwatch was sometimes difficult to read, especially the hundredths of seconds. This may have introduced a small margin of error in the measurements.

Depth Perception
The results of the depth perception tests showed significant differences in the effectiveness of the camera-equipped glasses and the endoscope in estimating depth in different scenarios. When using the camera-equipped glasses, participants experienced challenges with regard to accurately perceiving depth. The participants made significantly more errors and took longer to play the games with the VR glasses compared to without VR glasses.
The percentage differences in performance between using the VR glasses and not using them were substantial. For Dr. Bibber, the percentage difference in time taken was 176% and the difference in errors made was 62%. For the nerve spiral, the percentage difference in time taken was 161% and the difference in errors made was 359%. The program requirement was that the difference between using and not using the VR glasses should not exceed 50%. Clearly, this requirements was not met in any of the tests.

A key barrier that contributed to these challenges was the positioning and alignment of the cameras on the glasses. The wider distance between the cameras than between human eyes created a significant area where the images did not overlap, leading to double vision. This made it difficult to make accurate depth observations and led to situations where one hand blocked one of the cameras, further worsening stereoscopic vision and thus depth perception. Moreover, one of the participants reported that using the VR glasses was tiring on the eyes and sometimes led to nausea.
It is important to note that the games played are not fully representative of real-world application of camera-equipped glasses, which require users to passively watch and make decisions. Future research should focus on simulating realistic clinical scenarios to get a better idea of the performance and practicality of this technology.

In the endoscope method, participants were consistently able to correctly determine the order of beads in the cylinder. However, distance estimation proved difficult. It was not difficult to estimate how the beads were positioned relative to each other in space, but rather how far apart they were exactly. This could be because the test was conducted in an isolated environment where there were no reference points to help estimate distances. So adding reference points, such as a scale, could be an improvement. It could also be because it was not disclosed beforehand whether the beads were all the same size.

The percentage differences in distance estimation between the phantom viewed through the VR glasses and the actual distances were -37%, 13% and -16%. The program requirement was that the difference should not exceed 50%. This requirement was met.

Clinical Implications
The findings of this study can be placed in the clinical context. An important aspect is the resolution of the images. In the tests conducted, the endoscope performed significantly better than camera-equipped glasses. High resolution is important for performing surgery where detailed views of tissues and anatomical structures are essential for accurate decision-making. It can help surgeons detect abnormalities and perform precise procedures.
Colour reproduction of systems is important for clinical practice. It plays a major role in diagnostic accuracy, as it can help identify pathological tissues and other important details. Also, inaccurate colour reproduction can lead to misinterpretations which can result in errors in treatment or diagnosis.
In terms of depth perception, being able to accurately assess spatial orientation and distance to anatomical structures is very important for the person who is watching and possibly making decisions. However, the ability to perceive depth correctly is affected by the quality of the stereoscopic display. As observed in this study, challenges with camera-equipped glasses can hinder the accuracy of depth perception. Double vision can lead to confusion and inaccurate estimation of distances during operations.

Potential Improvements
To try to improve the glasses, various methods were tried to align the cameras solidly and accurately. These included using tyraps, blocks of wood, clips, tape and drilling holes for tyraps and a nail. Figure 7 shows one of the first versions of the camera-equipped glasses. After some modifications, the cameras were attached using tyre-wraps and a nail so the cameras do not slide up. A block of wood has also been used so that the cameras stay in the right position. These adjustments can be seen in figure 12b. However, the asymmetry of the glasses and manual work made perfect alignment difficult to achieve, which remains a challenge.
To overcome the problems of camera alignment and attachment, it is necessary to develop a more precise system for attaching and adjusting the cameras. One possible solution to this is to design a special frame that is adjustable at several points using bolts. For example, this could be done with a clinical trial frame. This is an adjustable pair of glasses used when fitting and testing different lenses during an eye test. The trial frame has the possibility of precise adjustments for the cameras to a rigid frame. Instead of using glasses, a mechanism similar to that of a headlamp could also be used to carry the cameras. This makes it possible to position the cameras closer together.

To improve the resolution of camera-equipped glasses, cameras with better quality could be used. Cameras of higher resolution and advanced image sensors can help capture sharper images. In addition, real time image processing software such as image stabilisation and noise reduction can help improve image quality. Using a dual lens system can improve the whole chain. These systems have an adjustable distance between the lenses to better mimic natural human binocular vision. This can help improve depth perception and spatial orientation.

Future Perspectives
In addition to the technologies already being used, the future of remote live 3D viewing in healthcare is expected to develop even further.
The metaverse is a concept that describes the merging of a physical and virtual world, accessible via computer. It combines 3D environments, real-time movement and sound, making you feel like you are in that virtual world. In healthcare, the metaverse can be used for medical training and education, virtual consultations between doctors and patients and remote collaboration between medical professionals. The metaverse is not a replacement for physical contact, but a tool to be used to improve care.
Augmented reality (AR) is a part of the metaverse. AR adds an extra layer of information to the physical world perceived, giving users an immersive virtual experience. AR allows virtual objects to be projected onto a real environment visible through a smartphone or smart glasses. In medicine, AR is mainly used in neurosurgery. In this specialisation, 2D images are used to plan surgeries and navigate during surgery. By projecting 3D images onto the surgical area, AR helps improve the outcomes of surgery by giving the surgeon a better view of the surgical area. [53]

Further studies could focus on developing a training program for medical staff using these technologies. This could be a standardised program to train both novice and experienced surgeons and healthcare staff to use these technologies during surgical procedures. Research could be conducted on how this program could contribute to improving surgical skills and quality of care, such as the impact on patient outcomes.

Another next step could focus on designing and developing a special frame on which the cameras can be securely and accurately attached to the glasses. This frame should provide adjustment options to optimise the distance and angle between the cameras. This could then be used to recreate a natural stereoscopic vision similar to human eyes. This could be done by 3D printing a frame to create a lightweight system.
An alternative approach could be to develop a new system where the cameras are not attached to glasses, but rather to a strap that can be worn around the surgeon's head. This system would offer flexibility in the positioning and adjustment of the cameras. The cameras could be positioned much closer together, avoiding double vision.

7 Conclusion

The investigation focused on the optimisation of stereoscopic imaging techniques for real-time remote viewing during surgical procedures and their impact on guidance, education and expert assistance in different applications.

The evaluation of stereoscopic imaging techniques showed both strengths and weaknesses for either camera-equipped glasses or the endoscope. The glasses caused minimal image distortion and low latency, but did not score well on resolution, colour reproduction and depth perception. The endoscope performed well on resolution, depth of field, image distortion and depth perception. However, colour representation was not accurate and latency was on the high side. In particular, the disappointing depth perception of the camera-equipped glasses highlights the need for further optimisation of this technology.

In the context of real-time remote viewing, image quality is very important for surgical supervision, educational purposes and expert assistance. Future research should focus on improving techniques and developing standardised training programs for medical staff.

Looking back at the objectives of this project, techniques were developed for capturing and transmitting stereoscopic images with an endoscope and camera-equipped glasses. Live streaming for 3D viewing was tested, for example through the latency test. The effectiveness of the systems was evaluated by several tests, such as depth perception and image distortion. These findings provide valuable insights for further optimisation.

Regarding the impact on guidance, education and expert assistance, it was found that differences in resolution between systems can affect the accuracy of visual information. The evaluation of depth perception and image distortion highlights the need for further improvement for optimal performance in surgical and educational settings.

References

[1] K. Stuij, "Deventer Ziekenhuis start met innovatieve VR-bril: 'Dit is een enorme winst voor de zorg'," *Destentor*, Jan. 2024.

[2] P. J. Choi, R. J. Oskouian, and R. S. Tubbs, "Telesurgery: Past, present, and future," *Cureus*, vol. 10, no. 5, p. e2716, 2018. 2168-8184 Choi, Paul J Oskouian, Rod J Tubbs, R Shane Journal Article Review United States 2018/08/07 Cureus. 2018 May 31;10(5):e2716. doi: 10.7759/cureus.2716.

[3] N. Raison, M. S. Khan, and B. Challacombe, "Telemedicine in surgery: What are the opportunities and hurdles to realising the potential?," *Curr Urol Rep*, vol. 16, no. 7, p. 43, 2015. 1534-6285 Raison, Nicholas Khan, Muhammad Shamim Challacombe, Ben Journal Article Review United States 2015/05/31 Curr Urol Rep. 2015 Jul;16(7):43. doi: 10.1007/s11934-015-0522-x.

[4] N. M. Weiss, A. Schneider, J. M. Hempel, F. C. Uecker, S. M. van Bonn, S. P. Schraven, S. Rettschlag, T. Schuldt, J. Müller, S. K. Plontke, and R. Mlynski, "Evaluating the didactic value of 3d visualization in otosurgery," *Eur Arch Otorhinolaryngol*, vol. 278, no. 4, pp. 1027–1033, 2021. 1434-4726 Weiss, Nora M Orcid: 0000-0002-0645-3591 Schneider, Armin Hempel, John M Uecker, Florian C van Bonn, Sara M Schraven, Sebastian P Rettschlag, Stefanie Schuldt, Tobias Müller, Joachim Plontke, Stefan K Mlynski, Robert Journal Article Germany 2020/07/03 Eur Arch Otorhinolaryngol. 2021 Apr;278(4):1027-1033. doi: 10.1007/s00405-020-06171-9. Epub 2020 Jul 1.

[5] L. Ganry, N. Sigaux, K. S. Ettinger, S. O. Salman, and R. P. Fernandes, "Modified gopro hero 6 and 7 for intraoperative surgical recording-transformation into a surgeon-perspective professional quality recording system," *J Oral Maxillofac Surg*, vol. 77, no. 8, pp. 1703.e1–1703.e6, 2019. 1531-5053 Ganry, Laurent Sigaux, Nicolas Ettinger, Kyle S Salman, Salam O Fernandes, Rui P Journal Article United States 2019/04/23 J Oral Maxillofac Surg. 2019 Aug;77(8):1703.e1-1703.e6. doi: 10.1016/j.joms.2019.03.026. Epub 2019 Mar 27.

[6] J. Dangen, Y. S. Hsueh, S. Y. C. Lau, S. Nagra, D. Watters, and G. D. Guest, "Live-streaming surgery during covid-19 using a 3d printed camera," *ANZ J Surg*, vol. 91, no. 6, pp. 1056–1058, 2021. 1445-2197 Dangen, Jordyn Orcid: 0000-0001-8674-6839 Hsueh, Yung-Hsin Sylvia Orcid: 0000-0003-4280-4376 Lau, Steve Yee Chiang Orcid: 0000-0003-2819-9704 Nagra, Sonal Watters, David Orcid: 0000-0002-5742-8417 Guest, Glenn Douglas Journal Article Australia 2021/06/15 ANZ J Surg. 2021 Jun;91(6):1056-1058. doi: 10.1111/ans.16797.

[7] B. N. Nasri, J. D. Mitchell, C. Jackson, K. Nakamoto, C. Guglielmi, and D. B. Jones, "Distractions in the operating room: a survey of the healthcare team," *Surg Endosc*, vol. 37, no. 3, pp. 2316–2325, 2023. 1432-2218 Nasri, Bao-Ngoc Orcid: 0000-0002-5561-2624 Mitchell, John D Jackson, Cullen Nakamoto, Keitaro Guglielmi, Charlotte Jones, Daniel B R01 EB005807/EB/NIBIB NIH HHS/United States Journal Article Germany 2022/09/08 Surg Endosc. 2023 Mar;37(3):2316-2325. doi: 10.1007/s00464-022-09553-8. Epub 2022 Sep 7.

[8] J. Majernik, "3d virtual projection and utilization of its outputs in education of human anatomy," *Multimedia Technology*, vol. 2, pp. 1–8, 03 2013.

[9] B. T. Backus, "Stereoscopic vision: what's the first step?," *Curr Biol*, vol. 10, no. 19, pp. R701–3, 2000. Backus, B T Journal Article Review England 2000/10/26 Curr Biol. 2000 Oct 5;10(19):R701-3. doi: 10.1016/s0960-9822(00)00707-7.

[10] Y. fu Chen, Y. Y. Lee, T. Chen, J. L. Semmlow, and T. L. Alvarez, "Review: Behaviors, models, and clinical applications of vergence eye movements," *Journal of Medical and Biological Engineering*, vol. 30, pp. 1–15, 2010.

[11] M. Giesel, A. Yakovleva, M. Bloj, A. Wade, A. Norcia, and J. Harris, "Relative contributions to vergence eye movements of two binocular cues for motion-in-depth," *Scientific Reports*, vol. 9, p. 17412, 11 2019.

[12] D. M. Levi, "Learning to see in depth," *Vision Res*, vol. 200, p. 108082, 2022. 1878-5646 Levi, Dennis M R01 EY020976/EY/NEI NIH HHS/United States R21 EY030609/EY/NEI NIH HHS/United States Journal Article Research Support, N.I.H., Extramural Review England 2022/07/17 Vision Res. 2022 Nov;200:108082. doi: 10.1016/j.visres.2022.108082. Epub 2022 Jul 14.

[13] K. W. Nam, J. Park, I. Y. Kim, and K. G. Kim, "Application of stereo-imaging technology to medical field," *Healthc Inform Res*, vol. 18, no. 3, pp. 158–63, 2012. 2093-369x Nam, Kyoung Won Park, Jeongyun Kim, In Young Kim, Kwang Gi Journal Article Korea (South) 2012/11/02 Healthc Inform Res. 2012 Sep;18(3):158-63. doi: 10.4258/hir.2012.18.3.158. Epub 2012 Sep 30.

[14] M. S. Saidon, H. Desa, P. Murugesa, S. Rosbi, and M. Rosbi, "Developing stereo vision system for object recognition and positioning of amad-r mobile robot,"

[15] A. W. Kusuma, "Design and analysis of streaming video media on 3g and 4g networks," *Jurnal Jartel: Jurnal Jaringan Telekomunikasi*, 2018.

[16] G. Su, Y.-C. Lai, A. Kwasinski, and H. Wang, "3d video communications: Challenges and opportunities," *International Journal of Communication Systems*, vol. 24, pp. 1261 – 1281, 10 2011.

[17] K. C. Celebi, S. K. T. Bailey, M. W. Burns, and K. Bansal, "Is virtual reality streaming ready for remote medical education? measuring latency of stereoscopic vr for telementoring," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 65, pp. 757 – 761, 2021.

[18] S. Singh, V. Sharma, P. Patel, G. Anuragi, and R. G. Sharma, "Telementoring: an overview and our preliminary experience in the setting up of a cost-effective telementoring facility," *Indian J Surg*, vol. 78, no. 1, pp. 70–3, 2016. 0973-9793 Singh, Suresh Sharma, Vivekanand Patel, Pinakin Anuragi, Gajendra Sharma, Raj Govind Journal Article India 2016/05/18 Indian J Surg. 2016 Feb;78(1):70-3. doi: 10.1007/s12262-015-1429-y. Epub 2016 Jan 11.

[19] K. C. Celebi, S. K. T. Bailey, M. W. Burns, and K. Bansal, "Is virtual reality streaming ready for remote medical education? measuring latency of stereoscopic vr for telementoring," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 65, no. 1, pp. 757–761, 2021.

[20] "Renkforce RF-VRG-200 3D-bril Zwart," June 2024. [Online; accessed 19. Jun. 2024].

[21] A. Mahboob, S. Husung, C. Weber, A. Liebal, and H. Krömker, "Smartphone as a stand-alone device for rendering, visualization and tracking for use during product development in virtual reality (vr)," *Volume 1: 39th Computers and Information in Engineering Conference*, 2019.

[22] S. Harada, T. Abe, J. Furumido, K. Takahashi, K. Hori, N. Abe, M. Kon, S. Murai, H. Miyata, H. Kikuchi, R. Matsumoto, T. Osawa, and N. Shinohara, "A prospective study of surgeons' workloads and associated factors in real-world practice," *Scientific Reports*, vol. 14, no. 1, p. 9741, 2024.

[23] "wetten.nl - Regeling - Wet aanvullende bepalingen verwerking persoonsgegevens in de zorg - BWBR0023864," May 2024. [Online; accessed 13. May 2024].

[24] "Welke landen horen bij de Europese Economische Ruimte (EER)?," May 2024. [Online; accessed 13. May 2024].

[25] "Handleiding Algemene Verordening Gegevensbescherming (AVG)," May 2024. [Online; accessed 13. May 2024].

[26] "Art. 5 GDPR – Principles relating to processing of personal data - General Data Protection Regulation (GDPR)," Oct. 2021. [Online; accessed 13. May 2024].

[27] "Hoe lang wordt mijn medisch dossier bewaard?," May 2024. [Online; accessed 13. May 2024].

[28] "NEN 7510: Informatiebeveiliging in de zorg - ICT in de zorg - Zorg & Welzijn," May 2024. [Online; accessed 13. May 2024].

[29] "Encryption - General Data Protection Regulation (GDPR)," Sept. 2023. [Online; accessed 13. May 2024].

[30] "Regulation - 2017/745 - EN - Medical Device Regulation - EUR-Lex," June 2024. [Online; accessed 19. Jun. 2024].

[31] H. Australia, "Endoscopy," *Healthdirect Australia*, Aug. 2022.

[32] R. Reilink, S. Stramigioli, and S. Misra, "Image-based flexible endoscope steering," pp. 2339–2344, 10 2010.

[33] "Otorhinolaryngology (ear, nose and throat (ENT) surgery)," Aug. 2023. [Online; accessed 29. Apr. 2024].

[34] L. Ma and B. Fei, "Comprehensive review of surgical microscopes: technology development and medical applications," *J Biomed Opt*, vol. 26, no. 1, 2021. 1560-2281 Ma, Ling Fei, Baowei R01 CA156775/CA/NCI NIH HHS/United States R01 CA204254/CA/NCI NIH HHS/United States R01 HL140325/HL/NHLBI NIH HHS/United States R21 CA231911/CA/NCI NIH HHS/United States Journal Article Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Review United States 2021/01/06 J Biomed Opt. 2021 Jan;26(1):010901. doi: 10.1117/1.JBO.26.1.010901.

[35] C. E. Marie Freebody, "Advances in Surgical Microscopes Pave the Way to Improved Outcomes," *Photonics Spectra*, Apr. 2017.

[36] D. Getty and P. Green, "Clinical applications for stereoscopic 3-d display," *Journal of The Society for Information Display - J SOC INF DISP*, vol. 15, 2007.

[37] D. K. Ho, "Using smartphone-delivered stereoscopic vision in microsurgery: a feasibility study," *Eye (Lond)*, vol. 33, no. 6, pp. 953–956, 2019. 1476-5454 Ho, Derek Kwun-Hong Orcid: 0000-0002-6209-3594 Journal Article Video-Audio Media England 2019/02/14 Eye (Lond). 2019 Jun;33(6):953-956. doi: 10.1038/s41433-019-0356-8. Epub 2019 Feb 12.

[38] Canon and T. Coleman, "3D VR with the dual fisheye lens," *Canon Europe*, May 2024.

[39] S. He and W. L. Davis, "Filling-in at the natural blind spot contributes to binocular rivalry," *Vision Research*, vol. 41, no. 7, pp. 835–840, 2001.

[40] R. Awati, "resolution," *WhatIs*, Aug. 2022.

[41] L. McKay, "Spatial resolution," *Radiopaedia*, Feb. 2024.

[42] "1951 USAF Resolution Calculator," June 2024. [Online; accessed 5. Jun. 2024].

[43] "USAF 1951 - determination resolulution," June 2024. [Online; accessed 5. Jun. 2024].

[44] S. Abasi, M. Amani Tehran, and M. D. Fairchild, "Distance metrics for very large color differences," *Color Research & Application*, vol. 45, no. 2, pp. 208–223, 2020.

[45] "File:Color Checker.pdf - Wikimedia Commons," May 2024. [Online; accessed 6. Jun. 2024].

[46] K. Thomsen, "A euclidean color space in high agreement with the cie94 color difference formula," *Color Research and Application*, vol. 25, pp. 64–65, 2000.

[47] S. Kuthirummal, H. Nagahara, C. Zhou, and S. K. Nayar, "Flexible depth of field photography," *IEEE Trans Pattern Anal Mach Intell*, vol. 33, no. 1, pp. 58–71, 2011. 1939-3539 Kuthirummal, Sujit Nagahara, Hajime Zhou, Changyin Nayar, Shree K Journal Article Research Support, U.S. Gov't, Non-P.H.S. United States 2010/11/23 IEEE Trans Pattern Anal Mach Intell. 2011 Jan;33(1):58-71. doi: 10.1109/T-PAMI.2010.66.

[48] M. Kleiber and C. Winkelholz, "Distortion of depth perception in virtual environments using stereoscopic displays: quantitative assessment and corrective measures," in *Stereoscopic Displays and Applications XIX* (A. J. Woods, N. S. Holliman, and J. O. Merritt, eds.), vol. 6803, p. 68030C, International Society for Optics and Photonics, SPIE, 2008.

[49] "Dokter Bibber online kopen?," June 2024. [Online; accessed 21. Jun. 2024].

[50] "Zenuwspiraal - Zenuwspel," June 2024. [Online; accessed 21. Jun. 2024].

[51] B. Lee, D. H. Cho, S.-C. Chung, G.-M. Eom, K.-S. Kim, and Y. J. Kang, "New optical design and image assessment of ultrahigh-resolution magnifying endoscope," *Japanese Journal of Applied Physics*, vol. 43, p. L1397, oct 2004.

[52] Q. Wang, A. Khanicheh, D. Leiner, D. Shafer, and J. Zobel, "Endoscope field of view measurement," *Biomed Opt Express*, vol. 8, no. 3, pp. 1441–1454, 2017. 2156-7085 Wang, Quanzeng Khanicheh, Azadeh Leiner, Dennis Shafer, David Zobel, Jurgen Journal Article United States 2017/07/01 Biomed Opt Express. 2017 Feb 7;8(3):1441-1454. doi: 10.1364/BOE.8.001441. eCollection 2017 Mar 1.

[53] I. Seddon, E. Rosenberg, and S. K. I. Houston, "Future of virtual education and telementoring," *Current Opinion in Ophthalmology*, vol. 34, no. 3, pp. 255–260, 2023.

Appendices

A Score Form

Vervorming van het beeld
Let op eventuele vervormingen in de positie, vorm of grootte van de ruitjes op het millimeterpapier, zoals krommingen.

Beoordelingsschaal:

- 1: Geen vervorming, alle ruitjes zijn gelijkmatig verdeeld en van uniforme grootte.
- 2: Zeer lichte vervorming, minimale afwijking in positie en grootte van enkele ruitjes.
- 3: Matige vervorming, duidelijke afwijking in positie en grootte van meerdere ruitjes.
- 4: Aanzienlijke vervorming, veel ruitjes vertonen afwijkingen.
- 5: Ernstige vervorming, alle ruitjes zijn afwijkend in positie en grootte.

Jouw score: X

B MATLAB Script

```
% Afbeelding inladen
image_color_test = 'liggend_endo_0021.jpg';
img = imread(image_color_test);

% Werkelijke RGB-waarden van de MacBeth ColorChecker
trueRGB = [
    115, 82, 68;
    194, 150, 130;
    98, 122, 157;
    87, 108, 67;
    133, 128, 177;
    103, 189, 170;
    214, 126, 44;
    80, 91, 166;
    193, 90, 99;
    94, 60, 108;
    157, 188, 64;
    224, 163, 46;
    56, 61, 150;
    70, 148, 73;
    175, 54, 60;
    231, 199, 31;
    187, 86, 149;
    8, 133, 161;
    243, 243, 242;
    200, 200, 200;
    160, 160, 160;
    122, 122, 121;
    85, 85, 85;
    52, 52, 52
];


% Selectie van de kleurpatches in de afbeelding (van links naar rechts, rij
```

```
% voor rij)
figure; imshow(img); title('Selecteer elk van de 24 kleurpatches (van links
    naar rechts, rij voor rij) en druk op Enter na elke selectie');

measuredRGB = zeros(24, 3);
for i = 1:24
    % Teken een veelhoek en wacht tot de gebruiker klaar is
    h = drawpolygon;
    wait(h);
    mask = createMask(h);

    % Bereken de gemiddelde RGB-waarden binnen de veelhoek
    redValue = img(:,:,1);
    greenValue = img(:,:,2);
    blueValue = img(:,:,3);

    measuredRGB(i, 1) = mean(redValue(mask));
    measuredRGB(i, 2) = mean(greenValue(mask));
    measuredRGB(i, 3) = mean(blueValue(mask));
end

close(gcf);

% Bereken het verschil tussen de werkelijke en gemeten RGB-waarden
colorDiff = sqrt(sum((measuredRGB - trueRGB) .^ 2, 2));

% Plot resultaten
figure;

% Plot werkelijke vs. gemeten RGB-waarden
subplot(1, 2, 1);
hold on;
for i = 1:24
    plot([1, 2], [trueRGB(i, 1), measuredRGB(i, 1)], '-r', 'LineWidth', 2);
    plot([1, 2], [trueRGB(i, 2), measuredRGB(i, 2)], '-g', 'LineWidth', 2);
    plot([1, 2], [trueRGB(i, 3), measuredRGB(i, 3)], '-b', 'LineWidth', 2);
end
xlim([0.5 2.5]);
xticks([1 2]);
xticklabels({'True RGB', 'Measured RGB'});
title('True vs. Measured RGB values');
ylabel('RGB Value');
hold off;

% Plot kleurverschil
subplot(1, 2, 2);
bar(colorDiff);
title('Color Difference (Euclidean Distance)');
ylabel('Difference');
xlabel('Patch Index');

% Zet de gemeten en werkelijke RGB-waarden voor elke patch in een tabel
patchIndex = (1:24)';
trueRGBStr = arrayfun(@(r,g,b) sprintf('[%3d, %3d, %3d]', r, g, b), trueRGB
    (:,1), trueRGB(:,2), trueRGB(:,3), 'UniformOutput', false);
measuredRGBStr = arrayfun(@(r,g,b) sprintf('[%3d, %3d, %3d]', r, g, b),
    measuredRGB(:,1), measuredRGB(:,2), measuredRGB(:,3), 'UniformOutput',
    false);
```

```
differenceStr = arrayfun(@(d) sprintf('%.2f', d), colorDiff, 'UniformOutput
    ', false);

resultTable = table(patchIndex, trueRGBStr, measuredRGBStr, differenceStr,
    ...
    'VariableNames', {'Patch', 'True RGB', 'Measured RGB', 'Difference'});

disp(resultTable);

% Bereken het gemiddelde verschil
meanDiff = mean(colorDiff);

% Bereken de standaarddeviatie van het verschil
stdDiff = std(colorDiff);

% Geef het gemiddelde verschil en de standaarddeviatie weer
disp(['Gemiddeld verschil: ', num2str(meanDiff)]);
disp(['Standaarddeviatie: ', num2str(stdDiff)]);
```

C Materials

Resolution Test

- Camera-equipped glasses
- Endoscope (Olympus OTV s300)
- Tripod
- Laptop with OBS Studio and TeamLink
- Smartphone with TeamLink
- 1951 USAF Resolution Test Chart
- Box to mount the test target
- Renkforce RF-VRG-200 3D glasses (VR headset)

Colour Accuracy Test

- Camera-equipped glasses
- Endoscope (Olympus OTV s300)
- Tripod
- Laptop with OBS Studio, TeamLink and MATLAB
- Smartphone with TeamLink
- MacBeth ColorChecker test chart
- Box to mount the test target

Depth of Field Test

- Endoscope (Olympus OTV s300)
- Tripod
- Laptop with OBS Studio and TeamLink
- Smartphone with TeamLink

- Hollow cylinder covered with millimetre paper
- Millimetre paper
- Renkforce RF-VRG-200 3D glasses (VR headset)

Image Distortion Test

- Camera-equipped glasses
- Endoscope (Olympus OTV s300)
- Tripod
- Laptop with OBS Studio and TeamLink
- Smartphone with TeamLink
- Millimetre paper
- Box to mount the test target
- Renkforce RF-VRG-200 3D glasses (VR headset)

Latency Test

- Camera-equipped glasses
- Endoscope (Olympus OTV s300)
- Tripod
- Laptop with OBS Studio and TeamLink
- Smartphone with TeamLink
- Smartphone
- Stopwatch

Depth Perception Test

- Camera-equipped glasses
- Endoscope (Olympus OTV s300)
- Tripod
- Laptop with OBS Studio and TeamLink
- Smartphone with TeamLink
- Hollow cylinder covered with paper and beads
- Games (Dr. Bibber and nerve spiral)
- Renkforce RF-VRG-200 3D glasses (VR headset)