
Otto-von-Guericke University
Faculty of Electrical Engineering and Information Technology
Chair for Healthcare Telematics and Medical Engineering



Thesis

for obtaining the academic degree
Master of Science (M.Sc.)

An Augmented Reality Tool for Resection Site Repair Phase during Laparoscopic Partial Nephrectomy

Author:

Seyedsina Razavizadeh

November 5th, 2020

First Examiner:

Prof. Dr. Christian Hansen

Second Examiner:

Dr. David Black

Supervisor:

Fabian Joeres, M.Sc.

Razavizadeh, Seyedsina:

Virtual Torchlight/Stethoscope:

*An Augmented Reality Tool for Resection Site Repair Phase during
Laparoscopic Partial Nephrectomy*

Master's Thesis, Otto-von-Guericke University

Magdeburg, 2020.

Abstract

Purpose Kidney cancer is the 13th most deadly diagnosed cancer. Surgical removal of cancerous tissues is established as an effective treatment, e.g. laparoscopic partial nephrectomy. While laparoscopic surgery has proven beneficial for the patients, the surgeons face more challenges and undergo more workload. Currently, no solutions are provided to aid the surgeons during the resection site repair phase (renorrhaphy). In this phase, sub-surface structures, e.g. renal vessels and collecting system, are at the risk of being punctured or damaged by the suturing needle.

Method An augmented reality solution, using visual and auditory displays is proposed, enabling the surgeons to see these structures. The visual display overlays a real-time Pseudo-chromadepth map at the tip of the instrument. The auditory display uses parameter mapping without reference for sonification of the required data, i.e. type, density, and distance of the risk structures.

Results The proposed solution was evaluated for the accuracy, duration, and workload improvements. The effects of the visual and auditory display, their combination, as well as their interactions, were investigated. Visual display shows 56.86% increase in accuracy ($p = 0.002$), 23.09% decrease in the workload rating ($p = 0.030$). The combination of both displays shows 42.21% increase in accuracy ($p = 0.007$). The auditory display shows improvement in the accuracy and workload, though not significant. The average task completion duration increased in all conditions.

Conclusion The results of this study suggest that augmented reality (specifically visual display significantly) can improve the workflow of resection site repair during laparoscopic partial nephrectomy and shows potential. Hence, further research and implementations of augmented reality methods in this workflow are encouraged.

Keywords Laparoscopic Partial Nephrectomy, Renorrhaphy, Augmented Reality, Visual Display, Visualisation, Auditory Display, Sonification.

Task of the Thesis in the Original:



FACULTY OF
ELECTRICAL ENGINEERING AND
INFORMATION TECHNOLOGY

Task Description of a Master Thesis

Seyedsina Razavizadeh

Student ID: 221317

Topic

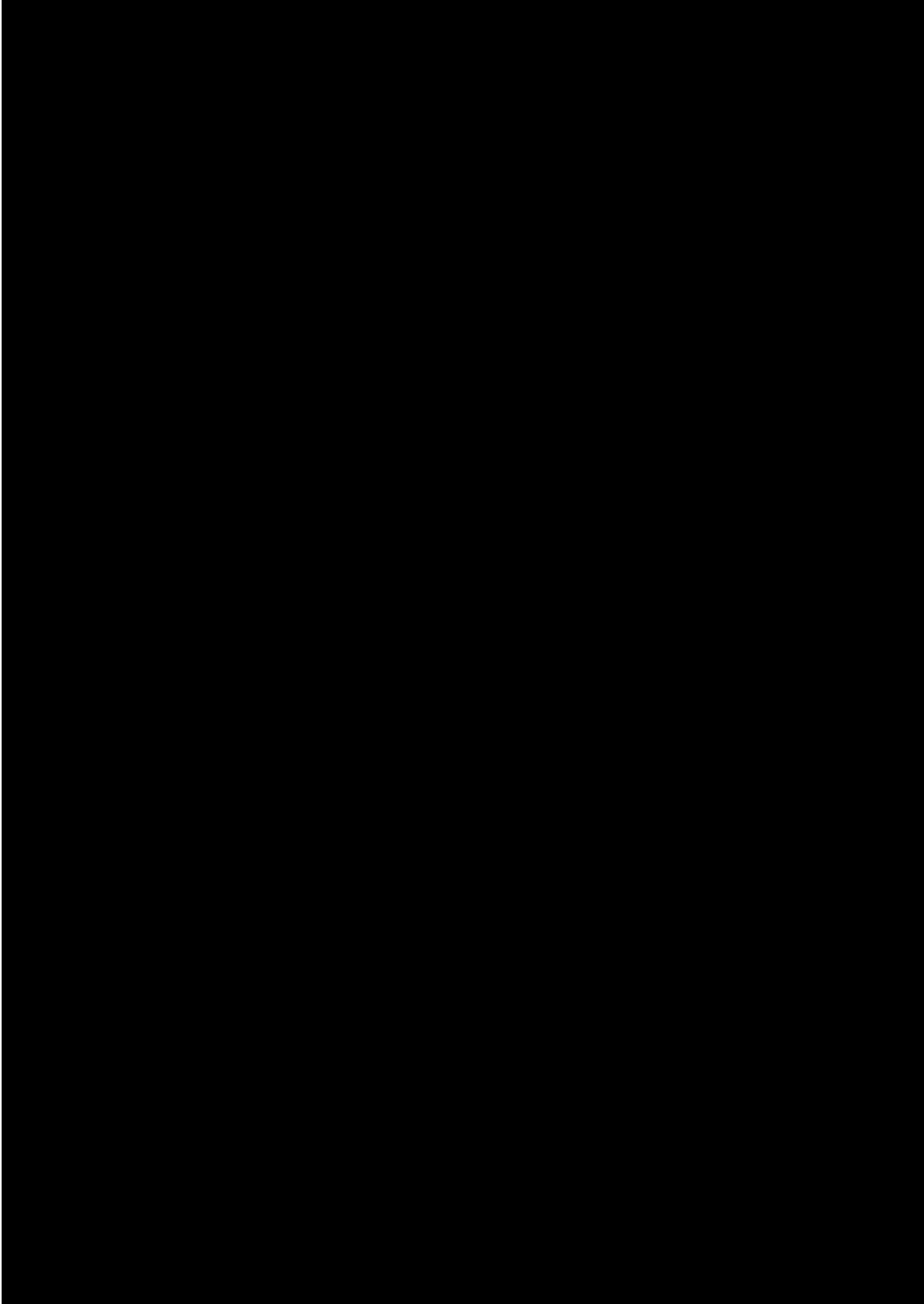
Virtual Torchlight/Virtual Stethoscope: An Augmented Reality Tool for Resection Site Repair Phase during Laparoscopic Partial Nephrectomy

Task Description

While Cancer is the second leading cause of death worldwide, if diagnosed in an early stage, treatment through surgical means could be a possible solution. One of these surgical procedures that have proven invaluable during the past years is Laparoscopic or Robot-assisted Resection (LR), e.g. Hepatectomy, Nephrectomy, etc. Through this minimally invasive procedure, in order to remove cancerous tissues, the organ will be partially or completely (radical) resected. Organ resection is a complex and challenging procedure with a high workload on the surgeon and somewhat not fully efficient workflow. In order to improve upon these shortcomings, different methods and software assistants have been proposed.

In the late years, with the rise of Augmented Reality (AR) technology, one can aim to better the workflow and reduce workload of resection procedure by relaying data and information through various means to the user. Most common AR approach is using Visual Displays to supplement information, e.g. overlay the camera feed with organ anatomy, risk structures, distance to certain structures, etc. Another promising AR approach, which has proven highly promising but still fairly far from commonplace in Computer-Assisted Surgeries, is using Auditory Display to convey a part of this information audibly, hence reducing the load of information from one sensory input and help with accuracy and workload of the procedure.

After resection of the cancerous tissues, the surgeon often tend to investigate the organ for risk structures that lie beneath the surface of resected volume. However since the resection volume is unknown, translation of this information from the image data (or software assistants) to organ is done by the surgeon which can prove challenging in many ways.



Declaration by the candidate

I hereby declare that this thesis is my own work and effort and that it has not been submitted anywhere for any award. Where other sources of information have been used, they have been marked, directly or indirectly.

The work has not been presented in the same or a similar form to any other testing authority and has not been made public.

I hereby also entitle a right of use (free of charge, not limited locally and for an indefinite period of time) that my thesis can be duplicated, saved and archived by the Otto-von-Guericke University of Magdeburg (OvGU) or any commissioned third party (e.g. *iParadigms Europe Limited*, provider of the plagiarism-detection service “Turnitin”) exclusively in order to check it for plagiarism and to optimize the appraisal of results.

Magdeburg, November 5th, 2020.

Razavizadeh, Seyedsina

Acknowledgements

First and foremost, I want to express my kindest gratitude to Professor Christian Hansen, who motivated me to work in the field of informatics, and gave me this opportunity to be a part of the Virtual and Augmented Reality research group, and work alongside its brilliant people. I would like to express my most profound appreciation to my supervisors, Dr. David Black and Fabian Joeres, who were there for me every step of the way, bearing through every question, providing guidance, and showing me the way. I would also like to thank Dr. Max-Philipp Stenner and Prof. Dr. Elena Azañón Gracia for showing me how to be precise yet open-minded and flexible when observing and analysing scientific phenomena.

I would like to thank my colleagues and friends; Sophia Freiwald for proofreading, participation, and continuous support and constructive feedback; Mohsen Askari, Idin D. Hoseini, and Onur Gülkökan for supporting me and providing constructive feedback during the work of this project; Ardiansyah Pringgo Esmondo for participation, and incredible vision in filming the final setup; Tonia Mielke for her exceptional help in making the evaluation setup possible, and participation; Daniel Reimert for providing an amazing segmentation of the kidney, and participation; and Mohammed Istiaque Amin, Ashik Aznad Anil, Vikram Apilla, Bankim Subhash Chander, Md. Amir Hossain, Elnaz Khosroshahi, Yohana Junitama Kustita, Yixue Lin, Charitha Omprakash, and Adithya Viswanathan for taking part in the studies without any expectations.

Contents

| | |
|---|-----------|
| Abstract | i |
| Declaration | iv |
| Acknowledgements | v |
| List of Figures | ix |
| List of Tables | x |
| List of Acronyms | xi |
| 1 Introduction | 1 |
| 1.1 Motivation | 1 |
| 1.2 Goals | 2 |
| 1.3 Structure | 2 |
| 2 Background | 4 |
| 2.1 Kidney | 5 |
| 2.1.1 Anatomy and Physiology | 5 |
| 2.1.2 Renal Cancer | 7 |
| 2.2 Laparoscopic Surgery | 7 |
| 2.2.1 Laparoscopic and Robotic-assisted Partial Nephrectomy | 8 |
| 2.2.1.1 Workflow | 8 |
| 2.2.1.2 Benefits and Limitations | 11 |
| 2.3 Augmented Reality | 12 |
| 2.3.1 Visual Display | 13 |
| 2.3.1.1 Optical See-through | 13 |
| 2.3.1.2 Projection-based | 15 |
| 2.3.1.3 Video See-through | 15 |
| 2.3.2 Auditory Display | 15 |
| 2.3.2.1 Parameter Mapping | 17 |
| 2.4 State-of-the-art | 18 |
| 2.4.1 Laparoscopic Partial Nephrectomy Software Assistance | 18 |
| 2.4.2 Augmented Reality Methods | 19 |
| 2.4.2.1 Visual Domain | 20 |
| 2.4.2.2 Auditory Domain | 23 |
| 2.5 Summary | 26 |

| | |
|---|-----------|
| 3 Method | 27 |
| 3.1 Requirements Analysis | 27 |
| 3.1.1 Functional and Non-functional Requirements | 27 |
| 3.1.1.1 Functional Requirements | 27 |
| 3.1.1.2 Non-functional Requirements | 28 |
| 3.2 Concept | 30 |
| 3.2.1 Core Features | 30 |
| 3.2.2 Visual Display | 31 |
| 3.2.2.1 Method One – Virtual Inspector | 31 |
| 3.2.2.2 Method Two – Real-time Pseudo-chromadepth Mapping | 33 |
| 3.2.2.3 Colour Selection Procedure | 34 |
| 3.2.3 Auditory Display | 38 |
| 3.2.3.1 Method One – Auditory Icon and Synthesised Tone | 40 |
| 3.2.3.2 Method Two – Musical Instruments with a Broad Tempo | 41 |
| 3.2.3.3 Method Three – Music Generation Cloud | 41 |
| 3.3 Implementation | 42 |
| 3.3.1 Development Environments | 42 |
| 3.3.2 Visual Displays | 44 |
| 3.3.2.1 Virtual Inspector | 44 |
| 3.3.2.2 Real-time Pseudo-chromadepth Mapping | 46 |
| 3.3.3 Auditory Displays | 48 |
| 3.3.3.1 Auditory Icon and Synthesised Tone | 48 |
| 3.3.3.2 Musical Instruments with a Broad Tempo | 50 |
| 3.3.3.3 Music Generation Cloud | 50 |
| 3.4 Discussion and Verdict | 51 |
| 4 Evaluation | 54 |
| 4.1 Concept and Design | 54 |
| 4.1.1 Accuracy | 57 |
| 4.1.2 Task Completion Time | 59 |
| 4.1.3 Workload | 60 |
| 4.1.4 Design Summary | 63 |
| 4.2 Setup and Settings | 63 |
| 4.3 Population | 66 |
| 4.4 Study Procedure | 66 |
| 4.5 Analysis and Results | 67 |
| 4.5.1 Descriptive Statistics | 67 |
| 4.5.2 Analysis of Variance | 69 |
| 4.6 Evaluation of the Requirements | 72 |
| 5 Discussion | 74 |
| 5.1 Outcome | 74 |

| | |
|---------------------------------------|------------|
| 5.2 Outlook and Future Work | 77 |
| 6 Summary and Conclusion | 80 |
| Bibliography | 82 |
| Appendix | 101 |

List of Figures

| | | |
|------|--|-----|
| 2.1 | The Anatomy of the Kidney. | 6 |
| 2.2 | Operating Room Setup. | 8 |
| 2.3 | Transperitoneal Port Placement and Patient Positioning. | 10 |
| 2.4 | Post-resection Renorrhaphy. | 10 |
| 2.5 | Mixed Reality Continuum. | 12 |
| 2.6 | Infrared Markers Used for Tracking. | 13 |
| 2.7 | Visual Display Methods in Surgical Literature. | 14 |
| 2.8 | Instances of AR Overlays. | 21 |
| 2.9 | The Virtual Window Proposed by <i>Bichlmeier et al.</i> | 22 |
| 2.10 | Chromadepth and Pseudo-chromadepth Implementations. | 23 |
| 2.11 | The Distance Mapping Method Proposed by <i>Schmidt.</i> | 23 |
| | | |
| 3.1 | Basic Comparison of Camera View Modes. | 33 |
| 3.2 | Two Gradient Methods. | 34 |
| 3.3 | Laparoscopic Images and the Applied Overlays. | 36 |
| 3.4 | Chi-Square Distance of the Overlaid Colours. | 36 |
| 3.5 | Chi-Square Distance in between the Overlaid Colours. | 37 |
| 3.6 | General Software Architecture of the Solution. | 43 |
| 3.7 | An Example of Pure Data (Pd) Environment. | 43 |
| 3.8 | Virtual Inspector in Use. | 45 |
| 3.9 | Pseudo-chromadepth Mapping Principle and Implementation. | 47 |
| 3.10 | Real-time Pseudo-chromadepth Mapping in Use. | 48 |
| | | |
| 4.1 | Target Point-clusters on the Kidney. | 58 |
| 4.2 | Definition of Distance Measurements in the Evaluation. | 59 |
| 4.3 | Evaluation Setup Overview. | 64 |
| 4.4 | Evaluation Setup Components. | 65 |
| 4.5 | Descriptive Statistics through Box Plots. | 68 |
| 4.6 | Descriptive Plots for RMA Results. | 71 |
| | | |
| A.1 | Adjustable Variables for Pseudo-chromadepth Mesh. | 101 |
| A.2 | NASA-TLX Questionnaire. | 102 |
| A.3 | Weighted NASA-TLX Pairs. | 103 |
| A.4 | Scatter Plots for the Evaluation Results. | 106 |

List of Tables

| | | |
|-----|---|-----|
| 2.1 | Concise Overview of LPN/RAPN Workflow. | 9 |
| 3.1 | Functional Requirements. | 28 |
| 3.2 | Non-functional Requirements. | 29 |
| 3.3 | Core Functions. | 30 |
| 3.4 | List of the Colours Used as an Overlay. | 35 |
| 3.5 | Discrete Auditory Parameters Categories. | 38 |
| 3.6 | Specifications of the Setup Computer. | 44 |
| 3.7 | Music Generation Cloud Base Frequencies. | 51 |
| 4.1 | Assistance Methods and AR Displays State. | 56 |
| 4.2 | RMA Results for within Subjects Effects. | 70 |
| 4.3 | Evaluation of Requirements. | 73 |
| A.1 | The Table of Descriptive Statistics. | 104 |
| A.2 | Evaluation Study's Assistance Methods and Points Progression. | 105 |
| A.3 | Correlation Coefficients and significance. | 106 |

List of Acronyms

| | |
|-----------------|---|
| ANOVA | Analysis of Variance |
| AA | Auditory Assistance |
| AD | Auditory Display |
| AR | Augmented Reality |
| CMYK | Cyan Magenta Yellow Black |
| CPU | Central Processing Unit |
| CT | Computed Tomography |
| EEG | Electroencephalography |
| GPS | Global Positioning System |
| GPU | Graphics Processing Unit |
| GUI | Graphical User Interface |
| HSL | Hue, Saturation, Lightness |
| HSV | Hue, Saturation, Value |
| LPN | Laparoscopic Partial Nephrectomy |
| NASA-TLX | National Aeronautics and Space Administration Task Load Index |
| NCI | National Cancer Institute (US) |
| NDI | Northern Digital Inc. |
| NHST | Null Hypothesis Significance Testing |
| NoA | No Assistance |
| OPN | Open Partial Nephrectomy |
| Pd | Pure Data |
| PM | Parameter Mapping |
| PN | Partial Nephrectomy |

| | |
|-------------|--|
| RAPN | Robot-assisted Partial Nephrectomy |
| RCC | Renal Cell Carcinoma |
| RGB | Red Green Blue |
| RGBA | Red Green Blue Alpha |
| RMA | Repeated Measures Analysis |
| RSRP | Resection Site Repair Phase |
| SD | Standard Deviation |
| SEM | Standard Error of the Mean |
| STL | Standard Tessellation Language |
| SWAT | Subjective Workload Assessment Technique |
| TLX | NASA Task Load Index |
| VA | Visual Assistance |
| VAA | Visual and Auditory Assistance |
| VAD | Visio-auditory Display |
| VD | Visual Display |
| WP | Workload Profile |

1 Introduction

1.1 Motivation

Ever since its discovery, cancer has had an enormous impact on annual death statistics [1–3]. As counteractive measures, scientists and medical experts have been investigating and introducing new methods to treat and ultimately cure cancer [4,5]. One of the more effective cancer treatments is the surgical resection of the cancerous tissue [5]. A category of these surgical approaches that has received an immense amount of attention in recent years is minimally invasive surgery. Laparoscopic surgery is a variant of the minimally invasive methods, in which the surgeon performs the surgery within the body using a camera and a set of surgical instruments [4]. This method is not limited to cancer therapy and can be used in various surgical scenarios.

Cancer is a global phenomenon, and the kidney is no exception. In 2018, renal cancer rated as the ninth most common cancer among men and 14th among women worldwide [6, 7]. It is a standard procedure to partially resect the kidney in order to remove the abnormality or cancerous tissue through laparoscopic partial nephrectomy (LPN) and its robotic counterpart robot-assisted partial nephrectomy (RAPN) [8]. These alternative surgical methods have the benefit of leaving a small footprint on the patient, compared to open surgery. The removal procedure is immediately followed by a resection site repair phase, where the surgeon applies sutures on the kidney to seal the wound and prevent further bleeding. Whether the task is carried out with or without a clamp, the necrosis of the kidney tissue and loss of blood is still of grave concern. The time-sensitive nature of the task would drive the surgeons to apply the sutures as fast as possible, increasing the probability of puncturing the internal risk structures of the kidney, i.e. vessels and urinary collecting system [9].

With the rise of Computer-assisted Surgery, many solutions have been introduced into the workflow of surgical procedures, including minimally invasive surgeries. Despite the ample instances of solutions and software assistants for laparoscopic surgeries, a solution for the reduction of error in this critical phase of the LPN/RAPN is yet to be proposed. The contents of the solutions mainly concern preoperative simulation and planning, and intraoperative navigation [9]. Commonly, in these software instances, a 3D visualisation of the organ is generated from medical image sets, and is investigated, manipulated, processed, and prepared before the surgical procedure. Some of the more recent solutions aim to bring the information to the exact site of the surgery. These techniques mainly depend on augmented reality (AR) methods to preview the content [10]. As a chief benefit, AR techniques grant access to the flow of information in situ, reducing the strain of using multiple information access points during a task, in our case surgery.

Given the fact that the surgeons approximate the location of the critical internal structure

based on the external landmarks on the kidney (e.g. hilus, upper and lower pole), and under time pressure nevertheless, providing the location of these structures can contribute immensely to the reduction of the possible errors, as well as temporal and mental demand of the tasks. In addition to this, exploiting AR techniques can further improve this impact mainly through reducing the time and workload of accessing multiple streams of information, due to the accessibility and availability of the information on site.

AR techniques are not limited to the visual domain. In actuality, AR methods have expanded across many sensory domains, i.e. visual, somatosensory, auditory, and olfactory [11]. Studies reports that multisensory information input, in comparison to mono-sensory input, can reduce sensory overload, have a favourable impact on information processing speed, and improve the efficiency of information perception [12]. Hence, in a sensitive scenario such as surgery, harnessing this effect can benefit the overall flow of the procedure. Two of the most immediate sensory systems that can be employed more unconstrained are visual and auditory senses. While vision provides fast and accurate acquisition of the data in a visually dominant task such as LPN, its limitations are inevitable, e.g. depth perception and line-of-sight. This shortcomings can be addressed with the help of an auditory feedback, which is omnidirectional, has an available auditory channel, and can be recognised quickly.

1.2 Goals

This thesis attempts to find a possible resolution to the following research question:

“How much can auditory and visual AR display methods contribute to the amelioration of the speed, accuracy, or workload during localisation of sub-surface structures of interest in laparoscopic resection site repair phase in partial nephrectomy?”

Thus, the possible effects of AR displays on duration, accuracy, and workload of LPN/RAPN’s resection site repair phase will be investigated.

To this end, an AR-based prototype is developed, which would supply an on-demand flow of spatial information. This information consists of the spatial situation of critical structures that lie beneath the resected surface, which is inspected by a surgical tool (e.g. a robotic arm or a surgical grasper). The AR methods used are of multisensory nature, with the final goal of reducing sensory overload, excessive unisensory dependency, mental pressure, and workload.

To conclude, the prototype will be evaluated and compared to the current conventions of the medical task. The effects and prospects of these AR displays will be analysed and discussed.

1.3 Structure

The chapters that are to be expected in this dissertation are outlined in the following.

- Chapter 2: Background

This chapter will provide the fundamental knowledge behind the main components of this

work, as well as reporting the state of the current technologies and research in the related fields.

- Chapter 3: Method

In this chapter, the primary process behind the realisation of the design will be discussed; This includes the requirements analysis, the concept behind the proposed methods, the design choices, and the final implementation route.

- Chapter 4: Evaluation

The main content of this chapter revolves around the evaluation and user study of the proposed system. The content consists of the evaluation's design, method and results.

- Chapter 5: Discussion

Here the final results and effects of the system will be discussed in detail. This chapter will cover the explanation behind the observed effects, the shortcomings, further works, and possible follow-ups.

- Chapter 6: Summary and Conclusion

As the final chapter, a summary and notable remarks of all aspects of the project will be presented.

2 Background

Since the recognition of cancer as a primary threat, scientists have been able to discover various treatments to tackle this phenomenon. According to the National Cancer Institute (NCI), cancer treatment modalities can be divided into the following groups [4],

- surgery,
- radiation therapy,
- chemotherapy,
- immunotherapy,
- targeted therapy,
- hormone therapy,
- stem cell transplant, and
- precision medicine.

With the implementation of these treatment methods, as well as improvements in the early detection of cancer, a steady increase in cancer survival rate has been observed in recent years [5].

Surgical removal of cancerous tissues has proven as an effective treatment [5]. In this treatment, depending on the cancer stage, the cancerous and its neighbouring tissues are partially or radically removed. Surgical approaches can be separated into two main groups, open surgery and minimally invasive surgery [4]. The procedure of open surgery is usually carried out with a large incision on the patient's body to provide direct access to the target tissue [4]. Alternatively, in minimally invasive surgery, a few small incisions are made, and with the help of surgical instruments (e.g. guiding catheters, graspers, laparoscopic camera, and robotic arms) the target tissue is reached, and the surgery is carried out [4].

Laparoscopic surgery is a subset of minimally invasive surgery during which a camera and a few trocars are inserted into the body through small incisions. This procedure enables the surgeon to operate within the body by manipulating these instruments.

With the recent advancement in computers and computational methods, a major addition to the workflow of surgery is a group of methods mainly known as Computer-assisted Surgery with a wide range of application from preoperative planning, intraoperative guidance, to post-operative diagnosis. These methods have been the focus of many research projects in the past decades, and have led to revolutionising the surgical procedure. Some of the more immediate examples of these methods are,

- image processing and analysis,
- digital image construction,
- visualisation of medical data,
- preoperative planning,
- surgical simulation,
- surgical navigation, and
- robotic surgery.

Delivery of information in these methods can be attained through different techniques. For instance, during the surgery, the information can be displayed and interacted with on a separate monitor, projected onto the patient's body, or on the active monitor (e.g. the visualised vessel tree that is overlaid onto laparoscopic video feed). These methods are more commonly known as augmented reality [13].

Augmented reality (AR) is the concept of supplementing the reality with additional (computer-generated) information on-demand or on-the-fly. This concept is not only limited to the visual enhancement, but can also be implemented via each of the sensory modalities (e.g. auditory, and haptic), and even multimodal approaches [11, 14].

A more recent achievement in minimally invasive surgery, specifically laparoscopic surgery, was the introduction of robotic assistants/surgical tools into its workflow [15, 16]. In this method, two or three robotic arms and a camera is inserted into the body, with the surgeon, manipulating the arms and the camera, having a more immersive experience and, somewhat, a more manageable workflow [15–17].

Within the scope of this dissertation, we mainly discuss the use of augmented reality in laparoscopic surgery, specifically the kidney. In the following subchapters, a brief background and literature review of the related works regarding the kidney, laparoscopic surgery, AR principles, AR visual displays, and AR auditory displays will be provided.

2.1 Kidney

2.1.1 Anatomy and Physiology

Lote delivers detailed insight on renal anatomy and physiology. Accordingly, kidneys are a vital part of our urinary and endocrine system. Located on both sides of the vertebral column and behind the peritoneum, these bean-shaped organs are accountable for a series of essential physiological functions,

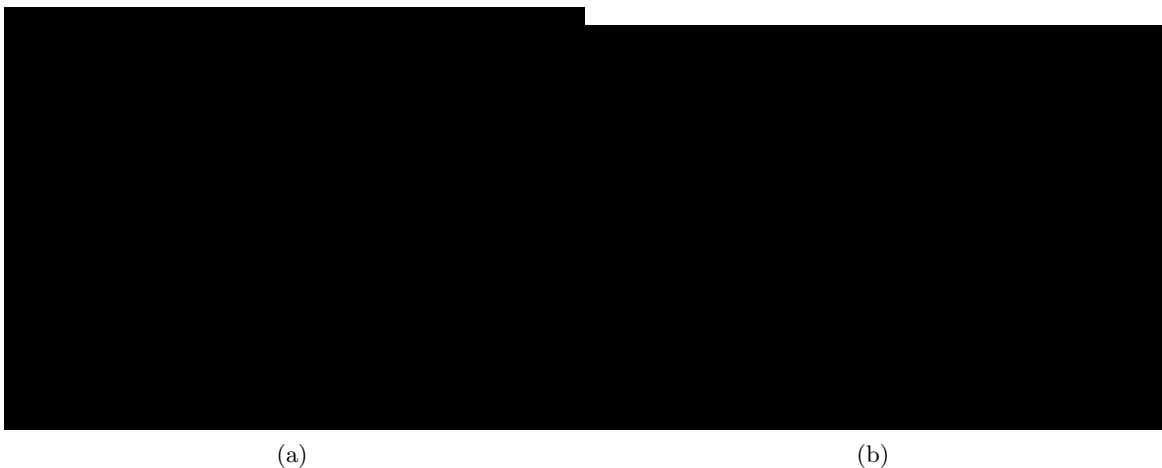
- acid-base homeostasis regulation,
- organism fluid supervision,

- regulation of electrolytes, e.g. sodium and potassium,
- toxin disposal,
- absorption of small molecules, such as glucose and amino acids,
- blood pressure management,
- hormonal production, and
- vitamin D activation.

Each kidney is approximately 12cm long, 6cm wide, and weighs about 150g [18].

At the medial section of the kidney, the renal hilus is located. The connections to the kidney are predominantly passing through this point, i.e. renal artery, vein, nerve, pelvis, and lymphatics. The kidney mainly consists of two regions, i.e. the cortex and medulla. The high blood input is filtered and processed through the nephrons, the working units of the kidney. The waste is mixed with water, forming urine. With medulla's supervision, the urine is guided through collecting ducts, calyces, and renal pelvis, where it follows its path through the ureter. A kidney produces 180 litres of filtrate per day. Besides urinary and filtration duties, the kidney is responsible for hormonal balance and secretion [18].

In addition to being filtered, renal blood flow, supply nutrients and oxygen to the kidney cells itself [18]. In case any damage falls on the renal collecting system or blood flow, the consequences would be grave and irreversible [18]. Hence, to reduce the risk of further complications, during the surgical procedure, unnecessary interactions with either of these systems are ill-advised. In accordance with the importance of these structures, the renal vessels and collecting system are identified as the main risk structures for the work of this thesis [9].



(a)

(b)

Figure 2.1: The Anatomy of the Kidney.
The cross-section of a) renal collecting system, and b) renal vascular system [18].

2.1.2 Renal Cancer

Renal cell carcinoma (RCC) is identified as the 13th most common cause of cancer death worldwide, with more than 140000 reported deaths (based on the most recent publication by the world health organisation in 2012) [19]. In 2018, the reported incidence rate of RCC was 5% in men and 3% in women of all diagnosed cancer cases [6,19].

To date, the surgical removal of the cancerous tissue remains the primary treatment method for non-metastatic RCC [8]. The surgical approach can be categorised into three main resection procedures [8],

- partial nephrectomy, retaining as much healthy renal tissue as possible to preserve renal function,
- nephrectomy, removal of the kidney, and
- radical nephrectomy, resection of kidney and the adrenal gland, surrounding tissue, and nearby lymph nodes.

Partial nephrectomy (PN) is the preferred method for stage I and II (according to TNM classification [20]) of RCC tumours [8].

2.2 Laparoscopic Surgery

United Kingdom National Health Service describes laparoscopic surgery. Accordingly, laparoscopic surgeries are mainly administered for abdominopelvic conditions and under general anaesthesia. As a subset of minimally invasive surgery, laparoscopic surgeries are carried out through a series of a small incision made on the patient's body at a required site. These incisions are used to insert surgical instruments, e.g. the camera, surgical scissors, and the gas tube. Carbon Dioxide gas is pumped through the tube to inflate the abdomen, enabling the surgeon to have more movement freedom and workspace. After the surgery, the gas is extracted, and finally, the incisions are stitched. [21]

The most common abdominopelvic fields where laparoscopic surgery is used are gynaecology, gastroenterology, hepatology, and urology. Compared to its more invasive counterpart, i.e. open surgery, laparoscopic surgery holds the following benefits for the patients,

- quick recovery time and short post-operation hospitalisation,
- reduction of haemorrhage and pain, and
- minimised scarring.

On the other hand, while this procedure holds more benefits for the patient, it is more taxing on the surgeon and requires a particular set of skills and training [22,23]. Unfortunately, the skills required for laparoscopic surgery are non-transferable, which suggests that the surgeons in the field of open surgery cannot switch in between these two methods easily [22,23].



Figure 2.2: Operating Room Setup.
 a) Laparoscopic Right Partial Nephrectomy [25], and b) Robot-assisted Left Partial Nephrectomy [17, 26].

2.2.1 Laparoscopic and Robotic-assisted Partial Nephrectomy

The first case of laparoscopic surgery was recorded in 1901, and over the past century, it has been implemented in many medical fields and treatments [15]. Owing to its success in these fields, the procedure was investigated for human renal conditions and was performed for the first time at 1990 in Washington University [15, 24]. With the advancement in robotic surgery, in 2004, the first RAPN was reported, and ever since then, it has achieved even more popularity among laparoscopic surgeons [15, 16, 24]. The examples of the operating room setup during LPN and RAPN is provided in figure 2.2.

2.2.1.1 Workflow

A survey by *Joeres et al.* Provides detailed information on the workflow of LPN/RAPN. A concise version of the conventional steps of LPN/RAPN can be found in table 2.1 [9, 17].

| | |
|----------------------------------|---|
| Preoperative Planning | Identification of size, location, proximity to risk structures, endophytic nature, and RENAL nephrometry score for the tumour based on medical imaging modalities, e.g. CT and MRI. |
| Preoperative Patient Preparation | Such as bowel preparation. |
| Patient Positioning | The patient is secured to a slight Trendelenburg positioned surgical table at an approximately 60° flank position (fig. 2.2 and 2.3). |

Continued

| Continued | |
|---------------------------------------|---|
| Port Placement | Based on the location of the tumour, required number of access ports for instruments, and surgeon's preference, a transperitoneal or retroperitoneal approach is selected, and then the incisions are made (e.g. fig. 2.3). |
| Insufflation and Instrument Insertion | Inflating the abdomen and inserting the instruments into the ports, e.g. camera, or robotic arms. |
| Operation Site Preparations | Mobilisation of the kidney (and the bowel, if necessary), and hilus dissection. |
| Investigation and Planning | Removal of excess tissue, e.g. fat, from resection site; the tumour, vessels, and risk structures examination and investigation (e.g. via ultrasound), and finally, marking the resection site. |
| Vascular Clamping | The necessary renal vessels are clamped, and the ischaemia timer is clocked. |
| Tumour Removal | The tumour is resected out of the kidney, resection site vessels are cauterised, and the tumour bed is sampled. |
| Resection Site Repair | Applying sutures to close major arteries and collecting system inlets, and repair the parenchyma. |
| Vascular Unclamping | The clamps are removed from the renal vessels to return blood circulation to the kidney. This step can also occur during the previous phase. |
| Tumour Extraction | The tumour is taken out of the patient's body. |
| Operation Conclusion | The surgical site is inspected to ensure haemostasis, the ports are extracted, the wounds are sutured, and the post-operative care is implemented. |

Table 2.1: Concise Overview of LPN/RAPN Workflow.

During the hilus dissection and intraoperative investigation, the surgeon tries to mark as many landmarks as possible. These landmarks are the cornerstone of the navigation and surgical procedure. By examining the laparoscopic video feed, one can locate the pronounced structures, e.g. ureter, renal artery, renal hilus, upper and lower poles. [9, 27–30]

To inspect the internal risk structures of the kidney, ultrasound and fluorescence is utilised. These methods can provide an insight into the vessels and the collecting system structures within the kidney, the position of the tumour, and possible undetected tumours. [9, 27–30]

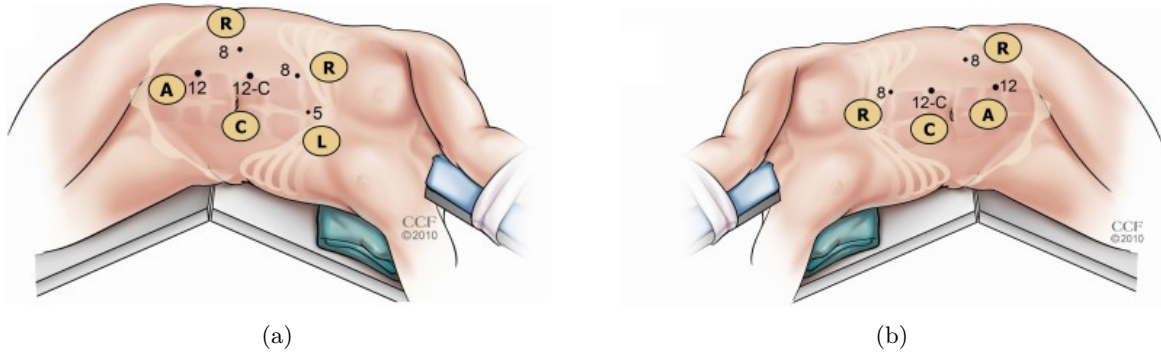


Figure 2.3: Transperitoneal Port Placement and Patient Positioning.

a) Robot-assisted Right Partial Nephrectomy, and b) Robot-assisted Left Partial Nephrectomy. $C = 12mm$ indicates camera port, $A = 12mm$ is the assistance port, $R = 8mm$ shows robotic instrument port, and $L = 5mm$ is for liver retraction [17].

After collecting all the necessary information, combined with previous anatomical knowledge, the surgeon recreates a mental image of the kidney and its contents. This mental image is a rough estimation of the kidney's actual internal structure. Maintaining this image for the entire duration of the surgery, specifically under stress and pressure of a surgical procedure, is mentally demanding and potentially highly error-prone. [9, 27–30]

During the resection site repair phase (RSRP), the surgeon uses this mental visualisation to apply sutures to the resection site. While applying the sutures (renorrhaphy), whether across or through the resection site (fig. 2.4), the surgeon should be mindful of vital structures that lie under the resection site [27–31]. These structures can be easily punctured or damaged by the suturing needle and cause intra- and post-operative minor and major complications [9]. Some of these complications are reported by *Turna et al.* [32]. Accordingly, post-operative complications can be separated into two main groups, i.e. urological and non-urological. Relative to the renorrhaphy task, mostly urological complications, e.g. post-operative haematuria, haemorrhage, and urine leakage, can be observed.



Figure 2.4: Post-resection Renorrhaphy.

a) through [28], and b) across the Resection Site [33].

It should be highlighted that this phase is executed under high temporal pressure due to ischaemia time. Additionally, the incision site and kidney landmarks are occluded by residual fluids and blood in the kidney [9]. Furthermore, kidney deformation, as a result of resection and the absence of blood circulation, affects the precision of this mental estimate [34–36].

2.2.1.2 Benefits and Limitations

More detailed comparisons of open partial nephrectomy (OPN) with LPN/RAPN have been presented in the literature [37–39]. A study by *Lucas et al.* shows that OPN offers shorter operation duration and (warm and cold) ischaemia time at the cost of higher blood loss and prolonged hospitalisation [37]. No cold ischaemia time is reported for LPN/RAPN in these results. Another study by *Park et al.* shows shorter ischaemia times for OPN and reduction in blood loss for LPN in RCC tumours with up to 1cm width [38].

Gill et al. conducted a high-volume study with 1800 cases of OPN and LPN with single renal tumours. The reported results show that LPN reduced operation time, blood loss, and the duration of post-operative hospital stay. On the other hand, LPN was also associated with extended ischaemia time, post-operative urological complications, and a higher number of required follow-up surgical procedures. The renal functionality and three-year cancer-specific survival in these cases were comparable. [39]

More recent research attempted to look into the advantages and disadvantages of LPN compared to RAPN [40–42]. In a study by *Jang et al.*, after investigating 127 cases of LPN and RAPN, results indicated no significant difference between LPN and RAPN except in surgical margin width [40]. This outcome indicates that with the implementation of RAPN, more healthy parenchyma can be preserved. Another set of results, provided by *Zhang et al.* from 1539 cases and 14 studies, suggests that similar operation duration, estimated blood loss rate, post-operative length of stay, intra- and post-operative complications, and positive surgical margin rate are observable in both approaches. In this study, warm ischaemia time has a significant improvement in RAPN [41]. On the other hand, *Choi et al.*'s research, consists of 2240 patients and 23 studies, shows that RAPN has a lower conversion rate to open or radical surgery, lesser variance in estimated glomerular filtration rate, reduces warm ischaemia time, and decreases the post-operative hospitalisation duration [42].

While the patients benefit from these improvements, surgeons, on the other hand, go through more mental and physical stress. Comparison studies between laparoscopic and open surgeries show that the surgeons suffer more mental and physical stress during laparoscopic surgery [43–46]; possibly, with more practice and experience, these strains could be reduced.

Some studies also investigated the workload of robot-assisted compared to conventional laparoscopic surgeries. The conjoint result of these studies suggests that under sufficient robotic instruction and training, RAPN can boost performance during tasks with high temporal demand [47, 48]. Theoretically, by improving the workflow of the laparoscopic surgery, one can reduce the surgeon's workload, and profit from the advantages that this method offers for the patients.

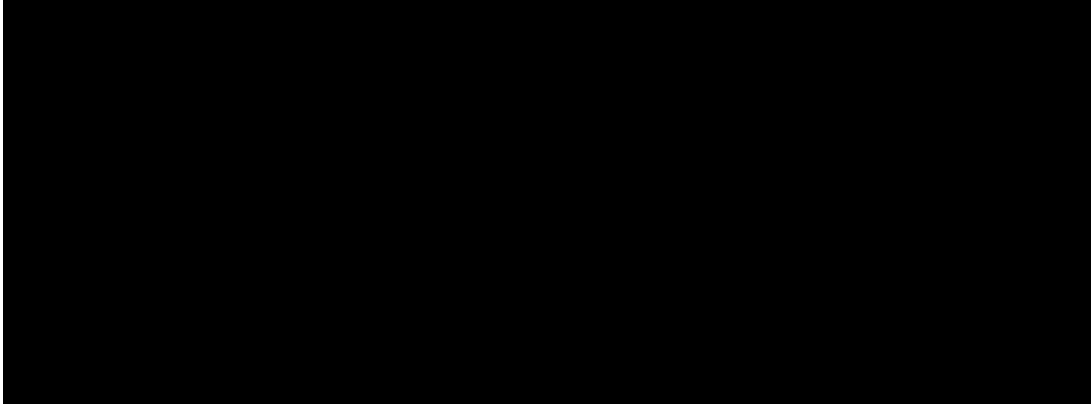


Figure 2.5: Mixed Reality Continuum.
Courtesy of *Anderson* [51].

2.3 Augmented Reality

In 1994 *Milgram and Kishino* defined the mixed reality continuum [49]. This continuum starts from reality and expands over to virtuality (fig 2.5). They also stated that mixed reality is the amalgamation of real and virtual worlds somewhere along this continuum, connecting real objects and environments with entirely virtual ones [11,49]. In 1997 *Azuma* established a widely accepted definition for AR. In this definition, any scenario on the continuum that can merge reality and virtuality contents, be interacted with in real-time, and is registered in 3D space, is considered an AR approach [50]. This definition disregards the sensory stereotypes of AR methods while placing more valid restraints on the definition, i.e. interactivity and registration [11].

To fulfil this definition of AR, registration and interactivity must be achieved. A standard solution for these requirements is the implementation of a tracking system through which registration (directly) and interactivity (directly, partially, or indirectly) can be realised. Tracking methods contain three main groups,

- stationary tracking, e.g. mechanical, ultrasonic, and electromagnetic tracking,
- mobile tracking, e.g. Global Positioning System (GPS), wireless networks, and inertia tracking, and
- optical tracking, i.e. marker-based, and markerless tracking.

As one of the most dominant physical tracking concepts, optical tracking is realised through vision-based analyses. In these methods, one or more cameras are used to investigate distinct features in the acquired images. These features can be either natural (e.g. shape edges) or artificial (i.e. markers). By recognition and classification of these attributes, the final tracking information is produced. [11]

As a subcategory of marker-based tracking, infrared tracking uses the infrared signal from active or passive markers to track the moving objects. The infrared signal is generated in the active markers via LEDs (fig. 2.6b). In the passive markers, this signal is cast by the camera



Figure 2.6: Infrared Markers Used for Tracking.
a) Retro-reflective Passive Markers, and b) Active Markers [52].

and reflected by the retro-reflective material of the markers (fig. 2.6a). The position of each marker is established through the time-of-flight from two (or more) cameras. The orientation of the tracked object, however, is estimated through spatial relations of a group of markers. To achieve six degrees of freedom in tracking space, theoretically, at the very least, three points are required, and practically, four points are required to formulate a unique solution. [11]

The infrared tracking offers an accurate spatial estimation of the trackers, that is mostly unphased by environmental interference. The emitted light provides rich information without any footprint on human's visual perception. Similar to other optical tracking methods, a clear line-of-sight between the camera and the object must be preserved. The attached sensors might prove cumbersome and inappropriate in some cases. Since active markers require electricity to function, they are often large (to carry batteries) or tethered via a cable to a power source. [11]

The AR methods are usually presented through AR displays. These displays differ from the traditional sense of the word, given that they are presenting virtual information in parallel to reality. Though AR methods are implemented through many sensory inputs, in the scope of this thesis, only visual and auditory methods will be further discussed.

2.3.1 Visual Display

As hinted by its name, a visual display (VD) aims to bring the virtual information to the visual domain. This goal is mostly achieved through three methods, i.e. optical see-through (fig 2.7c), projection-based (fig 2.7a), and video see-through (fig 2.7b) [11, 53, 54]. A significant amount of research and developments have focused on VD, making it the most common and known AR display [11].

2.3.1.1 Optical See-through

These displays mainly provide information by projecting them on a semi-transmissive and partially reflective optical object, e.g. lens or half-silvered mirror. The implementation of these displays

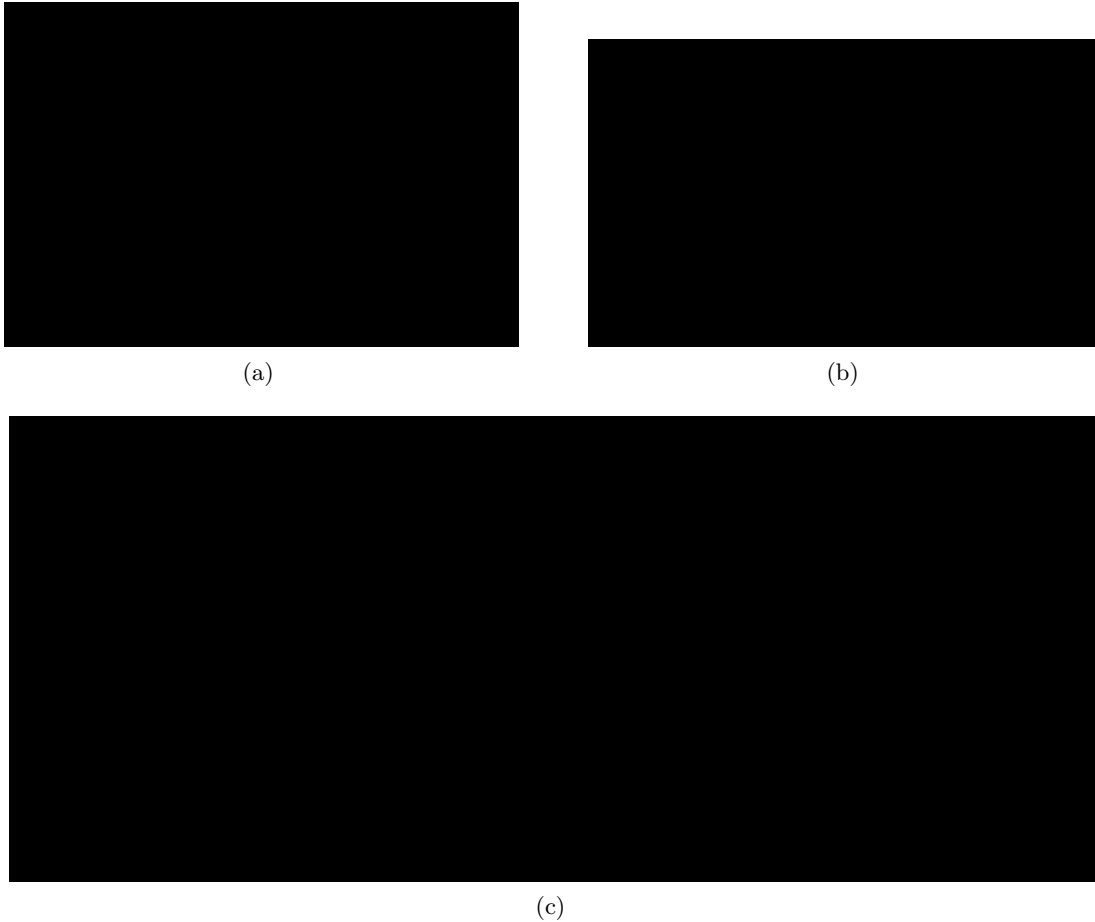


Figure 2.7: Visual Display Methods in Surgical Literature.

a) Projection-based [55], b) Video See-through [56], and Optical See-through with HoloLens [57].

can be seen in various hand-held, head-mounted, and static spatial setups (e.g. mirror) [11, 53, 54]. An example of a head-mounted optical see-through display is Microsoft HoloLens [11].

The advantages of this method are the presentation of the real-world contents in the actual resolution, without delays, and independent of power fails. Furthermore, the parallax that results from the camera positioning is not observed in this method. [11, 53, 54]

On the other hand, the virtual contents are usually rendered transparent as a result of the display elements (e.g. the lens) properties. The setup reduces the brightness and contrast of the reality, the field of view is limited, and shadows for depth perception cannot be rendered correctly. Special equipment is required for this setup that might not be compatible with every situation. While some of the required elements for this approach, e.g. mirrors, are more available and cheaper, the head-mounted displays can be quite expensive and sophisticated to implement and produce. [11, 53, 54]

2.3.1.2 Projection-based

In this display method, the contents of the virtual world are projected directly on the surface of the real object. This projection is commonly generated via light projectors in a hand-held, single, or multi-projector setup. [11, 53, 54]

Since no specific display medium is required, the user's field of view has no limitation. The virtual objects can be projected on any surface, whether simple or complex. The reality remains untouched in this method, which suggests that there is no input latency and resolution mismatch for the real-world content; however, this also restricts the mediation and manipulation capabilities of the virtual and physical contents. Furthermore, the virtual objects are only visible on the projection surface and not on the back-face of the object. The brightness, contrast, and the colours of the content, both virtual and real, can be affected by the environment and projection system. The virtual (and specifically dark) objects are rendered transparent, leading to colour misrepresentation, lack of virtual shadows, and possible loss of depth perception. As a final remark, it should be noted that projection-based approaches might require constant calibrations. [11, 53, 54]

2.3.1.3 Video See-through

To realise this method, a tracked camera and a monitor is utilised. The captured camera images are overlaid with the necessary information, resulting in a video, enhanced with virtual information. Both the real and virtual contents in this method are relying on the computer (e.g. smartphone), which is processing them. [11, 53, 54]

Since both the real and virtual objects are digitised, contents from both environments can be extensively manipulated and mediated. These objects can be easily synchronised, and the contrast, brightness, and colour of them can be corrected effortlessly. The virtual shadows and further shadings can be applied to improve depth perception. On the other hand, since the physical objects are rendered through the camera images, they are subject to latency and resolution limitations set by the display, computer, and camera. The field of view is also affected by the camera and the display (e.g. the monitor). [11, 53, 54]

2.3.2 Auditory Display

Similar to visual display, the environment can be augmented through auditory cues. This process is realised through sonification. As stated by *Barrass* and *Kramer et al.*, sonification is defined as "the use of non-speech audio to convey information and perceptual data" [58, 59]. This definition was later revised by *Hermann* as "the data-dependent generation of sound, if the transformation is systematic, objective and reproducible" [60]. Perhaps the most immediate instance of sonification is in the automotive industry. The first example is the car navigation system, in which the actions that the drivers must take to reach their destinations are conveyed through voice messages. The second example is the collision detection sensors in cars, where the repetition pace (rhythm) of the tone reports the distance to the object through parameter mapping [60, 61].

Besides voice messages and parameter mapping, audification and auditory icons (and Earcons) are the other approaches to auditory displays. Audification is the direct translation of physical information to the auditory domain. Stethoscope and ultrasound's auditory feedback are examples of audification. Auditory icons are familiar and natural recorded sounds that, directly or indirectly, represent an action or event in the environment. A well-known example of an auditory icon is the paper crumbling sound used when emptying the bin in computers operating systems. As a counterpart to the auditory icon, an Earcon achieves this presentation via a synthesised tone (e.g. a jingle). Auditory icons and Earcons are mostly short and brief. [60,61]

Based on the application, sonification can be separated into two main groups, spatial and non-spatial sonification. By exploiting human's spatial auditory perception and capabilities, one can deliver information regarding the position of objects through spatial sonification. This technique is mainly achieved through stereophony and spatialisation approaches. On the other hand, in non-spatial sonification, the information is relayed through perceptual properties of the sound. Some of these properties that are commonly used are pitch, rhythm (tempo), timbre, loudness, and fluctuation variability. [61]

A concise definition for the most common of these properties is provided in the following [60].

- *Loudness*, perhaps the most commonly known property of sound is the loudness. This property explains the amplitude of the sound wave. In more common literature this parameter is regarded to as sound volume or amplitude.
- *Pitch*, represents the fundamental frequency of the produced sound. The fundamental frequency is the lowest oscillation frequency in a sound wave. If the produced sound wave has a higher frequency, a higher pitch is perceived. The lower pitches are presented by bass frequencies, while the higher pitches are the results of treble frequencies. In the musical context, each musical note presents a fixed and predefined pitch. This pitch is mainly the fundamental frequency of a tone.
- *Rhythm*, this parameter is the regular speed at which a tone is played and repeated. This effect usually occurs due to introducing an onset into the sound wave. In the musical context, this onset is at the moment of creation of the note. While residual waves might still be present, the initial amplitude of the sound resembles a pulse onset.
- *Timbre*, this property is the main reason that we can distinguish between sound sources. The ability to distinguish between musical instruments, when they are producing the same pitch, is a result of timbre. Much is still unknown about this property, but it can be said that the variations in timbre are due to acoustic harmonics and their amplitudes when accompanying the fundamental frequency.

The characteristics of the target, whether it is dynamic or static, and the number of parameters to be mapped, are the factors that considerably affect the design choices for the mapping. Another factor to consider in the design is the availability of the sensory modalities. Whether

the auditory channel is fully available, partially engaged with other environmental information, the same information is provided in other sensory modalities, or other modalities are required to perform different tasks, are essential aspects to bear in mind while designing a sonification technique. [60,61]

2.3.2.1 Parameter Mapping

Parameter mapping (PM) relies on the human perceptual ability to distinguish the changes in frequency and time. The main principle behind this method is the mapping of different sets of information to different auditory parameters. PM for guidance can be mainly categorised into three groups. This classification is based on the effectiveness of the sound attribute in guidance efficiency and is as follows. [60,61]

The first PM class consists of methods that are *without reference*. In these strategies, the parameters are directly mapped to the psychoacoustic properties of the tones. The mapping is done by defining a minimum and maximum for the parameters value, establishing the function for the mapping (e.g. linear), and indicating the value at the target (whether it is maximum, minimum or in-between). The defined extrema are used to normalised the values that are to be mapped via the mapping function. In this method, the extrema and target tone are unknown to the user; hence the user is obliged to oscillate around the target to accomplish localisation. Alternatively, this scenario can be avoided by going through sufficient training. Ultimately, this method depends heavily on the perceptual resolution of the user. The accuracy of the task for each user is directly correlated to the Just Noticeable Difference in sound attributes. For instance, a trained musician has a higher resolution when discerning pitches than an untrained user. [61] [60]

Contrary to the previous method, the second class can be identified as approaches that are *with reference*. As the name suggests, a tone is generated that contains the correct parameters of the target. Parallel to this tone, another tone is produced, which is reporting the user's current spatial information. By analysing and comparing these two tones, the user can locate the target successfully. Similar to the previous method, the same extrema and mapping function should be provided in order to produce the tone which indicates the user's location; however, these requirements can be avoided through employing psychoacoustic parameters such as inharmonicity or roughness. The target sound can be any specific sound on the mapped spectrum. The benefit of these methods is that in comparison to previous methods the user is not required to oscillate and explore the sounds to locate the target correctly; however, in scenarios with no specific target or multiple simultaneous targets, these methods are not particularly effective. [61] [60]

The final category for PM methods comprises techniques that expand on the previous method with the addition of the *zooming effect*. In short, the zooming effect is the utilisation of a multi-scale variation for the mapping to increase the precision of parameter interpretation. As the user approaches the target, the mapping scale is altered, allowing a more accurate and quicker localisation. For instance, suppose a scenario where the low pitch is mapped to the

far distance and high pitch is mapped to the close distance. If the pitch variations in the far distance are occurring rapidly, in the close distance, these variations can happen at a much slower pace, notifying the user of even smaller changes in the pitch. Since this method is an expansion on strategies with reference, the absence, or simultaneous abundance of targets can prove unrepresentable. [61]

2.4 State-of-the-art

In order to achieve a better understanding of the literature and to address the objective of this thesis more efficiently, a brief review of the state-of-the-art is provided in this section.

2.4.1 Laparoscopic Partial Nephrectomy Software Assistance

A survey by *Joeres et al.* provides a detailed insight into the workflow, the required information, and the software assistants for LPN/RAPN [9]. According to this survey, the surgical challenges can be divided into three phases,

- Phase I, managing the renal vascular and hilar structures,
- Phase II, removal of the tumour, and
- Phase III, restoration of the renal defects, e.g. through renorrhaphy.

Based on the information needed by the surgeon in each one of these phases, specific software solutions have been investigated and proposed. These solutions aim to improve the workflow of the laparoscopic surgery by increasing the safety, effectiveness, and efficiency of the operation, or reducing the workload that is experienced by the surgeon. [9]

Phase I

The required information in this phase is regarding the anatomical structure of the kidney and clamping status. As discussed earlier, the renal vascular and collecting system are the risk structures during LPN/RAPN (section 2.1). Hence, providing information on renal structures, both internally and externally, before and during the surgical procedure, could significantly assist the surgeons. On the other hand, having a map of the renal substructures and the vessels increases the chances of making the right decisions and performing successful clamping. These structures are mainly connected to the tumour or are at the risk of being damaged during the surgery. [9]

With the aim of supplying the mentioned information, many methods have attempted to provide an anatomical and pathological 3D visualisation of the patient's kidney and its external connections [62–68]. These contents can be provided pre- or intraoperatively, through an AR method or on a secondary monitor.

Other approaches were proposed to assist the surgeon in localisation and recognition of vascular structures that are whether unrecognised in the preoperative images or are concealed during the

surgery [69–71]. These methods are based on the Doppler effect in ultrasound [69], fluoresces with near-infrared lights [70], and real-time pulsation detection [71].

The final group of these solutions aim to report the clamping status, and whether it has been successful or not [63, 69, 72–75]. In these strategies, the segmental (in selective clamping) [69] or whole (in full clamping) [72] perfusion of the kidney is investigated using Doppler effect, ultrasonic contrast agent [63, 73], or fluoresces [74, 75].

Phase II

The necessary information required for successful tumour localisation and extraction can be relayed through various 3D visualisations of the patient’s intra- and extra-renal structures. These methods aim to provide these contents during the operation [68, 76–82]. The contents can be generated before or during the operation and presented by AR or an alternative monitor.

As an alternative, information regarding the tumour can be acquired intraoperatively through ultrasound or contrast agent imaging [83–85]. These approaches use ultrasonic image overlay [86, 87], fluorescence imaging [70, 88], or tumour detection algorithms for ultrasound images [89].

The next group of methods attempt to provide tumour-specific information, such as the tumour and its supplying vasculature visualisation, safety margins, contours, and segmentation uncertainty [78, 82, 90–94]. These methods are mostly achieved through 3D visualisation of the content through AR.

The final group investigate the appropriateness of the tumour excision. One of these methods uses contrast agents such as fluorescence, to identify the remaining tumour tissues [88]. The other method uses a saline-filled bag and ultrasound probes to estimate the correctness of negative resection margins [95].

Phase III

The survey does not report any related work in this phase. This lack of previous works might be due to the uncertainty of the information caused by factors such as kidney deformation, and unknown resection site surface and volume. The lack of solutions for this challenging phase of the surgical procedure is a gap in the current literature which is investigated in this dissertation.

2.4.2 Augmented Reality Methods

Many studies proposed AR solutions for the workflow of laparoscopic surgery. A survey by *Bernhardt et al.* attempts to provide an overview of these strategies [10]. This survey also points out various benefits offered by the AR methods in these surgical procedures, which are mentioned in the following.

The guidance offered by AR enables constant identification of sub-surface and risk structures. This information can be used to prevent unnecessary damage to the risk structures and enhance the decision-making process of the surgeons. By providing resection trajectories and planes,

as well as safety margins, one can increase the chance of sparing as much healthy tissue as possible, while achieving negative margins. Furthermore, these contents can increase the focus and efficiency of the surgeons by relieving them from mentally connecting and recreating the information from multiple sources. Due to the small field of view in laparoscopic cameras, some information ambiguities can occur during the surgery. AR methods can clear these uncertainties by providing insight into them, boosting the surgeon’s spatial awareness. [10]

AR methods also present some challenges and limitations. *Bernhardt et al.* identified some of the shortcomings and challenges in AR. While most of the literature focuses on more immediate challenges, such as on-the-fly processing and interactivity, registration precision, and robustness, a portion of the studies address the perceptual limitation of AR. Two main limitations have been recognised by the literature, i.e. depth perception and content presentation. [10]

While designing the delivery method used for presenting the data, one must always consider three vital questions, i.e. “what to present”, “when to present it”, and “how to present it” [96]. The importance of considering these questions is thoroughly discussed in the work of *Dixon et al.*, where the possibilities of the surgeons being distracted from the real environment due to inappropriate placement of the AR content is argued [56]. Furthermore, the presentation of the content without reporting uncertainty and errors can lead to attention tunnelling. Under this effect, the surgeon disregards the actual data in favour of the virtual data, believing that the virtual data are presenting every necessary information, with absolute accuracy and precision [56]. This issue can be solved by presenting the uncertainty and errors in the visualisation. [10]

As pointed out by *Bogdanova et al.* and *Bernhardt et al.*, the depth perception is a significant factor in AR designs. In the absence of depth cues, human’s visual perception significantly fails in estimating the spatial relations of the objects. Occlusion and motion parallax can be identified as the primary depth cues [10,97]. By correctly occluding the posterior objects, a robust depth cue can be presented to the user. Minor depth cues, such as stereoscopy, shading, shadows, and the perspective view can improve the depth perception even further. Hence, providing at least one of these visual cues is highly advised. In addition to visual cues, *Bork et al.*, *Choi et al.*, and *Plazak et al.* suggested auditory cues to improve depth perception [98–100].

With the mentioned benefits and limitations and the potential of AR methods, many researchers are encouraged to invest in AR solutions for laparoscopic surgeries. In the following subsections, the related works in the field of visual and auditory AR will be reported.

2.4.2.1 Visual Domain

An AR implementation of anatomical information regarding the kidney and its relevant structures can be found in a number of studies [79–82]. *Chen et al.* proposed a method for LPN in which a virtual model of the kidney is constructed from the preoperative CT images. The renal tumour, kidney cortex, extra- and intra-renal arteries and veins are included in this visualisation. During the surgery, the model is registered manually on the kidney and viewed on a secondary monitor. The registration is then corrected using the data from an intraoperative ultrasound probe. The



Figure 2.8: Instances of AR Overlays.
Implementation by a) *Chen et al.*, and b) *Sengiku et al.*

final product of this registration is shown in figure 2.8a. The goal of this method is to control negative margins and enable image-guided tumour resection through an AR assisted method. [79]

Sengiku et al. proposed a similar method for RAPN in which a manual registration accompanied by an optical flow algorithm determine the orientation and position of the 3D generated content (fig 2.8b). The 3D volumes are constructed from preoperative CT images and consist of kidney cortex, arteries, veins, tumours, and urinary tract. During the evaluation surgery, the visualisation was available on the operating room’s monitor as well as the surgeon’s viewport. The method aims to improve tumour localisation and resection lines. [80]

Various AR visualisation methods for surgical procedures have been proposed over the years, to improve the efficiency and the impact of the information delivery [101–103]. Even though these strategies are not specific to LPN/RAPN, they are valid and applicable in this field as well.

Bogdanova et al. established the importance of depth perception during Minimally invasive surgery and its current limitation [97]. To solve these issues, various solutions were proposed, e.g. stereoscopy, shadows, Chromadepth, edge enhancement, light and contrast scattering, Virtual Window, and kinetic depth [97, 103, 104].

Bichlmeier et al. proposed an AR video see-through method, i.e. the Virtual Window, to address the depth perception challenges for in-situ medical visualisations [104]. In this method, the content is superimposed on the actual body of the patient, exploiting the motion parallax and occlusion to convey depth cues. Virtual Window was evaluated alongside six other visualisations of the spinal column, where it was recognised as one of the best visualisation methods for relaying depth cues. The examples of this visualisation can be found in figure 2.9.

In 1987 *Steenblik* introduced a novel depth encoding method for presentation of stereoscopic images, known as *chromadepth* [105]. In this method, the depth information is encoded into the entire visible colour spectrum. These colours are band into a smooth gradient, representing the

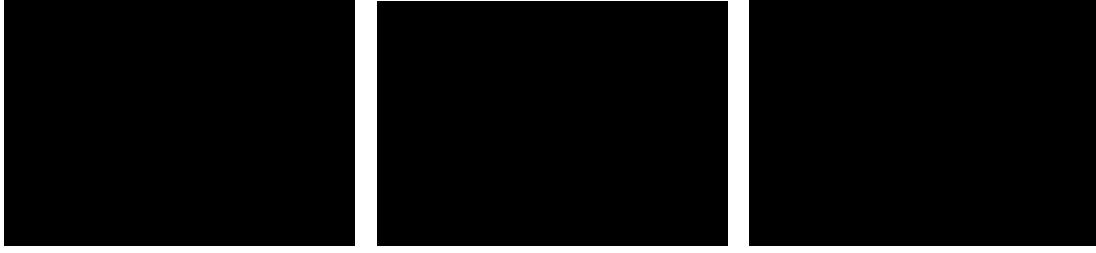


Figure 2.9: The Virtual Window Proposed by *Bichlmeier et al.*

distance of the structure from the camera or a certain point. Often the gradient starts from red as closest, followed by orange, yellow, green, and ending with blue as the furthest (fig. 2.10a). The main benefits of this method are,

- depth information is available in a single 2D image,
- the image can be interpreted even with unaided eyes, and
- independent of the user’s head position, contrary to 3D glasses.

The Chromadepth was later on implemented in vessel visualisation, to improve depth assessment and interpretation of angiography data [101, 102]. *Ropinski et al.* explored and expanded the idea by introducing *Pseudo-chromadepth*, where colours are mapped to a selective band rather than the whole spectrum (fig. 2.10b) [101]. This selectiveness is especially promising in the surgical environment where visualisation colours are constrained. *Joshi et al.* implemented the method as well, to establish a vesselness enhancement visualisation (fig. 2.10c) [102]. According to these two studies, when it comes to understating the depth cues, the Pseudo-chromadepth is highly preferable to the other methods [101, 102].

Further examples of Pseudo-chromadepth can be found in the work of *Kersten-Oertel et al.* and *Lawonn et al.* where this approach was compared to other depth perception enhancement methods (fig. 2.10d and 2.10e) [103, 106]. *Kersten-Oertel et al.* reported that Pseudo-chromadepth and aerial perspective (which is referred to as *fog* in the study) had the best outcome. In the study by *Lawonn et al.*, while the proposed technique showed the best result, the Pseudo-chromadepth method showed promising outcome as well.

Another approach was proposed by *Schmidt*, which exploits Chromadepth. In this scenario, the Chromadepth is used in reverse to the previous examples of this method. The Chromadepth is applied to a mesh which changes colours according to the distance from the vessel. The colour-coding resembles a distance-coded heat map. This method was designed for assistance in needle navigation during laparoscopic liver surgery. The technique was compared with other strategies, where it was recognised as the best solution for avoiding the risk structures. The distance mapping principle is depicted in figure 2.11. [107]

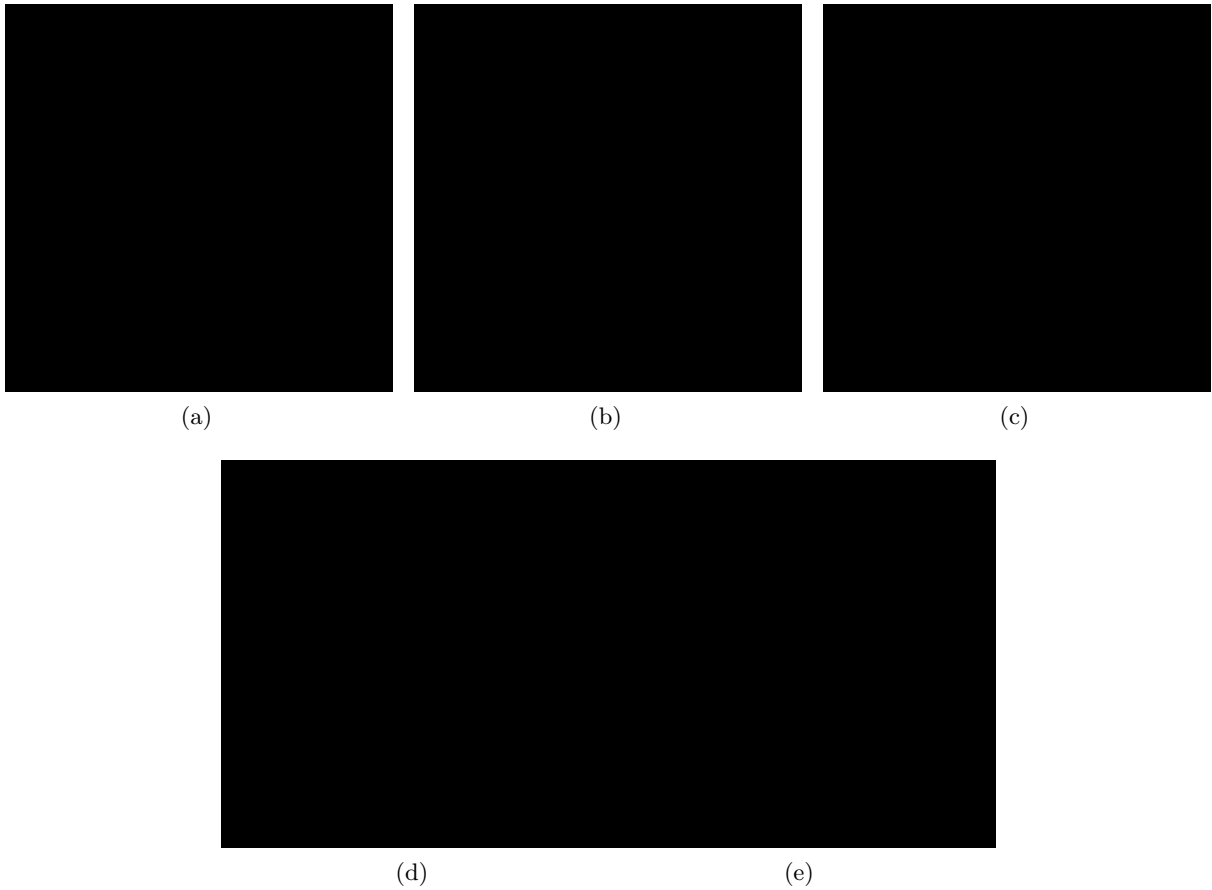


Figure 2.10: Chromadepth and Pseudo-chromadepth Implementations.
 a) Chormadepth by *Ropinski et al.*, and Pseudo-chromadepth by b) *Ropinski et al.*, c) *Joshi et al.*, d) *Kersten-Oertel et al.*, and e) *Lawonn et al.*

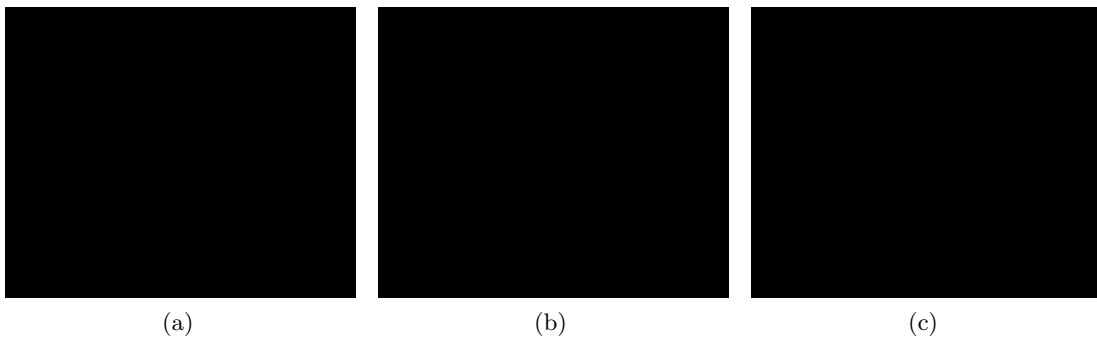


Figure 2.11: The Distance Mapping Method Proposed by *Schmidt*.
 Showing the instrument a) far from, b) near, and c) close to the structure.

2.4.2.2 Auditory Domain

A survey by *Black et al.* investigated auditory AR implementations in minimally invasive surgery [12]. This survey included a total of 13 approaches [98, 108–119]. The sonification methods used in these approaches are Earcons (and auditory icons) [108–115], and parameter mapping [98, 116–119]. This survey established three main motivations that can be observed in

the literature;

- increasing spatial awareness regarding the nearby anatomical and critical structures (e.g. surrounding the needle tip),
- reducing the focus shift between the monitor and the surgical site, and
- improving navigational data interpretation.

Additional motivations mentioned in the literature for implementation of AD are receiving information on visually occluded objects, reducing memory load and workload, decreasing simulation sickness and vertigo, and replacing the tactile feedback, which is unavailable during the robotic surgery. [12]

Furthermore, this survey revealed some merits of the sonification in this field. These benefits can be summarised as,

- auditory cues are omnidirectional and are free of any line-of-sight limitations [116,117],
- compared to visualisation, generation of tones are inexpensive [116,117],
- compared to the vision, the auditory channel is relatively open [116,117],
- auditory perceptual capability in detecting multiple audio streams [116,117], and
- a faster recognition of the events in high-stress tasks (e.g. using alarms) [59].

By implementing sonification in surgical tasks, these studies generally observed enhanced spatial recognition of the target structures [111–113], increased placement precision [98,111,118], a higher similarity in the resection volume [109], enhanced orientation [112], lesser workload demands [115], and decreased complications rate [111]. However, the prolonged duration of the task [111,118], and a higher amount of resected healthy tissue during volumetric resection [109] were some of the disadvantages of these techniques. [12]

The discussions in these studies pointed out a series of factors to consider while designing an auditory system. Sound and noise are a common occurring in the operating room; as shown by [120,121], it is much higher than the recommended levels (up to $93dB$). The level and sources of these sounds should be identified and considered during the design of the auditory display [98,108,109,113,115,116,118]. Another essential design parameter is ensuring that the tones are only generated when they are required [115]. By taking these two factors into account, one can ensure that the generated tones are successfully heard, do not fade in the background noise, nor overload the auditory channel of the user. Furthermore, customisable tones that can address the user's preference, and more natural tones (e.g. musical instruments) can be effective design choices [118]. [12]

Woerdeman et al. investigated the effects of auditory feedback on volumetric resection. In this study, the auditory display was compared with standard image-guided surgery and an

image-injection method. The system produces a soft warning tone when the resection tool reaches $5mm$ distance to the tumour's inner outline. The tone continues rising in volume and pitch until the outline is reached. If the tool passes the outline, a constant pure-tone is generated, alarming the surgeon. Even though no significant effect on the quality of the resection was observed, surgeons subjectively reported an improvement in the workflow of the procedure. [110]

Hansen et al. proposed an intraoperative guidance system for open liver surgery, where the resection path information is provided through auditory feedback. In this method, the resection line was presented as the target. The distance between the resection line (on both sides) and the instrument tip is categorised into *safe*, *warning*, and *outside* zones. While in the safe zone, a tone is generated, indicating the correct path of the resection tool. The tone in the warning zone aims to guide the user towards the safety margin. By mapping the distance to the rhythm and pitch of the tone, the system attempts to achieve this goal. As the tool drifts further away from the resection path, the rhythm and pitch of the tone are decreased. The pitch is representing both the distance and side of the resection line. In comparison to visual guidance alone, the combination of the visual and auditory assistants improved the accuracy and ergonomics (i.e. changing view between the monitor and situs) of the workflow, while increasing the duration of the task. [118]

Black et al. used two sonification methods, i.e. with and without reference, to assist needle insertion. These methods were carried out in three sequences, tip placement, alignment of the needle, and needle insertion. The first method generates two pulses sequentially, the first tone reports the user location, and the second tone presents the target location. The x-axis information is mapped to the rhythm of the tone, while the y-axis is presented by the pitch. In the alternative method, the rhythm indicates the x-axis value. When the user reaches a certain safe distance from the centre line of the x-axis, the tone changes to a higher pitch, but the rhythm mapping stays similar to the x-axis mapping. The results of the study showed a favourable outcome for the precision and confidence of the surgeons while using the auditory feedback. [119]

Bork et al. aimed to improve the depth perception of visual medical information by introducing audio to a purely visual workflow. In this method, a spherical propagation is generated from the instrument tip, accompanied by a metronome sound. When this sphere collides with a structure of interest, a secondary tone is played, and the structure is immediately rendered. The results of the study showed significant improvement in the precision of the task when AD was present. However, similar to other studies, the duration of the task is longer when any AR assistance is provided (with statistical significance). [98]

Black et al. proposed a stand-alone auditory support system for needle placement to reduce the surgeon's dependence on vision. This method could also be used in combination with VD. Spatial sonification was implemented in this method. The y-axis values are mapped to the pitch of the generated tone, ranging from $130Hz$ (below the target) to $523Hz$ (above the target). The x-axis is directly mapped to the position of the tone, i.e. whether it is heard in the left or right ear. The method was evaluated for accuracy, time, and workload, where a significant improvement in

workload was observed. This study also reported a significant increase in the completion time of the task under AD. However, no significant effect on the accuracy was found. [122]

A study by *Plazak et al.* aimed to assess and validate the workload experienced during an image-guided neurosurgery procedure while using an auditory display. The sonification method in this approach is a combination of a pure-tone and white noise. As the user approaches the target, the white noise would gradually fade, until no white noise is present. The method was designed and studied in a 3D localisation task. During the task, an EEG sample of the user was recorded to provide cognitive load measurements. Afterwards, a subjective workload test was carried out as well. The final results showed increased accuracy in the presence of AD compared to simple VD. The EEG cognitive load measurement was significantly higher in AD when compared to a combination of VD and AD. The subjective workload values were significantly increased compared to only VD or combination of VD and AD. [99]

2.5 Summary

While beneficial for the patient, laparoscopic surgery has always proven to be a more complex procedure for the surgeons in comparison to their alternative, open surgery, and LPN/RAPN are no exceptions. The AR displays have been implemented in the workflow of LPN/RAPN with the common goal of reducing the surgeon's workload and increasing the safety and accuracy of the procedure. The most common AR display used in these methods is VD. However, based on the recent research, AD also seems like a potential candidate for achieving this goal, or improving the workflow further by accompanying VD.

A study of the literature revealed that no current solution is provided for the RSRP. During this phase, the localisation and recognition of critical structures within surgical area falls entirely on surgeon's anatomical knowledge, ability to interpret spatial information, and recreating a 3D mental image from ultrasound images and other sources. This complicated process becomes even more arduous under the intense time-sensitivity of this phase of the procedure.

We believe that by providing the surgeon with spatial information of the current critical structures within the requested surgical area, the workflow of this surgical phase can be ameliorated. An AR approach can ease the access to this information and help the surgeon by reducing the duration of this phase, increasing the accuracy of localisation and recognition of structures, and in turn, decrease the workload necessary for accomplishing the task. These effects, ultimately, lead to fewer complications during the suture administration, as well as post-operative and long-term renal deficiencies.

3 Method

In this chapter, the methodology of the project is discussed in detail. The chapter begins with a report on requirements analysis, followed by a thorough description of the concepts. Afterwards, the implementation techniques for the concept is discussed. Finally, the outlook of each implemented method is analysed, and the final decisions are reported in the discussion and verdict section.

3.1 Requirements Analysis

As the first step in designing the solution, one should draw the outlines of the project to have a clear description and plan for the upcoming design steps. In order to achieve its goal, our solution must provide the user with a real-time interactive AR approach. This solution must also help the user with recognition and localisation of the renal structures through an accurate registration. In addition to these conditions, the system should be user-friendly, performant, robust, and practical. Since this work is a preliminary prototype of the described solution, more complex factors such as patient motion, organ deformation, and cavitory resection site is not investigated.

3.1.1 Functional and Non-functional Requirements

As a part of this preparation, analysis of functional and non-functional requirements is always a crucial stage. These requirements can be identified through interviews with users or in the literature of the similar field. This work mainly rely on the literature to determine these requirements (works done by [9, 10, 12, 107, 123]). In the following sections, these requirements and the result of the analysis will be discussed.

3.1.1.1 Functional Requirements

The functional requirements are every set of tasks that a system must perform in order to achieve its goal. These requirements are the major elements of a system; in other words, without them, the system cannot function.

Based on our research question, one can outline the critical design concerns for the solution of this work. The most dominant of these concerns is the realisation of our solution through AR. According to the definition of AR established in section 2.3, to properly implement AR, interactivity and registration must be addressed. Furthermore, the importance of interactivity, (accurate) content registration, and their recognition as two of the main challenges in medical AR systems are thoroughly discussed in the work of *Bernhardt et al.* [10]. Accordingly, component

| | |
|------------------------------|--|
| Augmented Reality | Based on the goal and research question of this project, the information must be provided using an AR method. |
| Component Tracking | To achieve interactivity and registration, the movement of all the relative objects, such as the world anchor and instruments, must be continuously tracked, whether with an external camera or an internal sensor (e.g. robotic joint (mechanical) tracking). |
| Registration | Based on the definition of AR and its requirements in medical domain, the virtual model must be registered with the minimum possible error. |
| Recognition and Localisation | Based on the required information reported by the surgeons, the recognition and localisation of the structures must be accomplished with the minimum possible error. |

Table 3.1: Functional Requirements.

tracking, as a mean for interactivity and registration, must be implemented to ensure fulfilment of the AR definition, as well as addressing the challenges indicated by this survey. The required information by the surgeon must be generated and interacted with in real-time. Additionally, an accurate registration of the virtual content on to the real-world objects is required [10].

Joeres et al. established the required information for this phase [9]. This information consists of sub-surface structures at the resection site and their spatial relations. During RSRP, surgeons require this information to be able to recognise and locate these structures in order to reduce the risk of damaging them. Our research question also remarks on this requirement. As our second most crucial design concern, our system must be able to address this requirement. By providing this information, the solution will assist the surgeon in localisation and recognition of the sub-surface structures. An overview of the functional requirements is presented in table 3.1.

3.1.1.2 Non-functional Requirements

As mentioned earlier, besides functional requirements, a system has sets of requirements that are defined as non-functional requirements. These requirements should be performed to achieve the set goals, and are not equally vital in the functionality of the system as their counterpart. In our case, these requirements can be separated into three categories, as mentioned in table 3.2, i.e. usability, performance, and presentation.

After investigating the **usability** factor for our system, it can be split into three subcategories. The first subcategory, *satisfaction*, is the friendliness and accessibility to the features of the system, that the user has experienced. The next topic focuses on the *efficiency* of the system, indicating the effort put into accomplishing the task using the proposed approach. Finally, the *effectiveness* of the system investigates the presence of improvements while using the system. This categorisation is based on ISO 9241-11:2018, where a thorough description of each one of these requirements is provided [124].

| Usability | |
|--------------------|--|
| Satisfaction | Based on user’s experience with the system. |
| Efficiency | Improvements in the workload and duration of the task with the help of the system. |
| Effectiveness | Improvement in the accuracy of the task with the help of the system. |
| Performance | |
| Display Frame Rate | The speed of which the content is updated in the displays. |
| Latency | Intra- and inter-component latency of the whole system. |
| Robustness | Resistance to internal failures and errors, as well as erroneous inputs. |
| Presentation | |
| Minimalistic | Contrasting, yet non-distracting, and blendable presentation of the content. |
| Colour-coding | Use of conventional colour sets which are suitable for realistic environments. |
| Auditory Cues | Use of distinguishable auditory cues for an intuitive and seamless experience. |

Table 3.2: Non-functional Requirements.

As mention in the context of *Bernhardt et al.*’s survey, the **performance** of the system comprises the *display frame rate*, *robustness*, and *latency* of the system [10]. Both displays should provide a standard output relevant to the human’s visual and auditory perception. In the case of visuals, a frame rate lower than 12 images per second is perceived as a series of still images [125]. On this account, a sense of motion is experienced when the frame rate is higher than 12. However, more comfortable experience is achieved in a frame rate higher than 24 [10,126], which should be the minimum requirement of this concept. A notable aspect of the system is latency. The latency could refer to both intra- and inter-component latency [107]. This response time can result in delayed or missing information, as well as disruption of real-time interactivity. The latency requirement explores the possibilities of such occurrences. The robustness investigates the resistance of the system towards internal and external errors, and possible failures, e.g. system crashes, unavailable tracking input, drastic content misplacement and registration errors, and sizable amounts of jitters and noise [10,107].

To have a more intuitive and practical interface, one should consider the **presentation** of the content to the user [10,12]. While this concept is mentioned as “relevance of visualisation” by *Bernhardt et al.*, it can also be amended to incorporate sonification as well. The content should be presented in a *minimalistic* way such that it does not interfere with the natural workflow of the task [10,12,96]. The information should be presented in a way which has the necessary contrast with the environment. The presentation should also maintain an acceptable level of

situational integrability without distracting the user from the primary task. Another factor in the presentation is the intuitiveness of the design, which in this case can be discussed as the familiar *colour-coding* and *auditory cues* [10, 12, 107]. The visualisation and sonification should be represented through elements that are well-established for the user in order to have a more coherent impact.

It should be noted that the tracking and content registration and its respective improvements and errors were not in the scope of this thesis. The registration topic is discussed in detail in another thesis, which was an ongoing project in parallel to this work [127].

3.2 Concept

3.2.1 Core Features

The essence of this project can be encapsulated as,

“An AR tool that provides localised spatial information on kidney’s internal risk structures, through an inspection tool, e.g. a robotic arm or a surgical grasper; in other words, one can observe what lies precisely in front of the instrument, just like a torchlight or stethoscope.”

Following the requirements mentioned in section 3.1, one should define a set of core features for the designs, in order to remain faithful to the project’s objective and retain an acceptable level of comparability among them. These core functions, found in table 3.3, must be observed in both VD and AD.

A crucial design aspect in the content presentation is the minimalistic way of information delivery. Even though having ample content could aid in keeping a complete view, the excess amount of information could lead to distraction, frustration, and confusion. Hence, presenting the necessary content, when it is needed, is an important design factor. This notion holds even more true in the AR setups since the information is overlaid directly on real-world elements and might even be occluding vital information. In the case of RSRP, this aspect can be even more crucial, given the immense gravity of the surgical site visibility and time-sensitivity of the procedure. To demonstrate, suppose the complete visualisation of the internal risk structures of the kidney is

| | |
|-----------------------|---|
| Localised Information | The generated content should be exclusive to the requested inspection zone. |
| Type Indication | Must display a clear distinction between vessels and urinary structures. |
| Spatial Data | The spatial information such as dimensions, distance from the instruments, and spatial relations between structures must be reported. |

Table 3.3: Core Functions.

overlaid on the VD. In this situation, perceptual and interactive issues (e.g. in depth and spatial perception) are the probable outcomes. Even with the correction of these perceptual challenges through proper visualisation methods, the complexity of investigating a whole structure in order to find local information could still prove overwhelming, let alone under time pressure. This problem is even more emphasised in the auditory realm since only two streams of auditory input can be processed simultaneously [128]. If one were to hear every possible detail of the current model, it would only appear as a distracting noise. This sound clutter would simply add up to the frustration of the task and detract the focus of the user. Accordingly, providing the portion of the information, which is requested by the user, could diminish many of these complications. In this way, the user could prepare for the content, focus on the upcoming information, and have a more efficient uptake of the presented content.

As discussed in section 2.1, the targeted risk structures (for the work of this thesis) are the renal vascular and collecting systems. The distinction between these structures must be effortless. The user must be able to recognise each of these structures intuitively. For this purpose, specific properties should be assigned to each type. These properties can be defined through different visualisation and sonification approaches, e.g. shading and colours for the VD, and auditory icons and Earcons for the AD. The intra-sensory contrast between the presentation of these structures must be sufficient; so that the user would predominantly bypass the doubt or uncertainty in the recognition process.

As a principle factor in our design, spatial information plays a critical role. It is based on this information that the surgeon would decide whether to apply the suture at the current place or not. The distance in which the structures lie, recognition of superficial structures, and their relative positions to their neighbouring structures, are all the requirements that this factor should present. In addition to spatial coordinates, the size of the structure is also of interest. The size of the vessels or the urinary ducts could be a possible criterion for the selection of critical structures; hence it impacts the suturing decision made by the surgeon. Supplying these sets of information to the user across all AR displays can provide a multisensory input, which in turn would decrease the sensory overload.

In the following subsections, the details of each proposed design are discussed.

3.2.2 Visual Display

The VD serves the purpose of delivering the generated information to the user's visual domain. To this end, two VD methods were conceptualised, which are explained in detail in the following subchapters.

3.2.2.1 Method One – Virtual Inspector

The underlying feature of this display is the ability to stream a video of the virtual sub-surface structures. The user would move the inspection instrument on the resection surface, and similar

to a camera, the instrument would show a real-time image of the internal risk structures. The method resembles the functionality of a magnifying glass when scanning a text and an Otoscope.

Following the virtual content's registration at the surgical site, the tracked instrument is attached with a virtual camera at its tip. This camera follows the instrument's movement, enabling the user to inspect the virtual objects that are present directly in front of the camera. In this system, the type of each structure can be easily colour-coded or shaded using standard visualisation methods.

The occlusion of valuable on-screen information must be avoided at all cost. Hence, the stream from the virtual camera should either be cast on a secondary monitor or be provided minimally in the same visual output as the laparoscopic video stream. As mentioned earlier in section 3.1, information delivery must be as minimalistic as possible. Hence, having the content on the same display as the video stream would prevent the physical demand of looking at two monitors, as well as averting the eyes from the surgical site. Furthermore, in RAPN, where the surgeon uses a control unit with a dedicated 3D display, requiring the user to switch between this display and a secondary monitor is an insensible choice.

The placement of virtual content on the main screen should be on a case-by-case basis. In respect to the entry point and the surgical site, the content should always occupy the bare minimum of the screen. The generated video stream can be cast on a movable window, giving the user the ability to change the location of it. A preferable location for the stream is near the viewport's edges.

The virtual camera can use two standard viewing modes, namely perspective or orthographic mode (fig. 3.1). The orthographic view has been implemented in many medical solutions, specifically imaging techniques, e.g. radiology. This method provides a wholesome view of the scene, which often comes with the cost of depth perception (fig. 3.1 right). Furthermore, the size of an object in the image is distance-irrelevant and is always reported as exact proportions (relative to the image size). On the other hand, in perspective mode, the camera gives more information regarding the spatial relations between the viewed objects. The size and the position of the objects in the image are relative to their distance from (and distortion of) the camera (fig. 3.1 left). As the depth perception has been indicated as a primary challenge in AR, any means, which could ameliorate this situation, should be exploited. Hence, the perspective mode is selected as the default camera view mode of this concept, where the orthographic view would remain as a selectable option.

Although the Perspective view can relay the relative spatial information of the objects in the scene, estimating the actual distance from the camera is an error-prone task. As a plausible solution, the distance between the surface mesh and the virtual camera could be measured and reported. Using this approach could contribute to the reduction of inaccuracy in distance estimation. It should be noted that the point from which the distance is measured must also be displayed.

3.2.2.2 Method Two – Real-time Pseudo-chromadepth Mapping

The alternative approach to the virtual inspector is based on the Pseudo-chromadepth principle, instead of a direct camera view, to display the internal structures. One of the major features of Chromadepth is that, in addition to two-dimensional information, one can embed depth information using colour-coding conventions. The instances of these methods were mentioned in section 2.4.2.1.

Given the time-sensitive nature of the task, the rapid flow of comprehensible information is crucial. Using colours or brightness can hasten the interpretation of this information drastically. If these colours were to represent the distance of the object from a certain point, one could create a colour-coded image of the objects that lie in front of the point. Based on our object's surface mesh, given decently minute distance calculation points, an image of the object can be recreated as a real-time Pseudo-chromadepth map.

The gradients can be applied in two ways, either continuous or posterised (colour-banding). By defining a distinct border between colour steps, posterisation improves the classification of the distances. However, as shown in figure 3.2, this may have an impact on the depth and spatial interpretation. On the other hand, a continuous gradient can provide a better depth perception while impacting the distance estimation. A compromise can be made through the creation of a highly varying gradient, which provides the user with relative information through a higher fluctuation while keeping comprehensible depth information.

In light of these arguments, choosing the proper colour set becomes even more essential in this design. As mentioned before, the colour sets must be distinct from the background environment, in our case organs, glands, fat tissue, and blood. Furthermore, each structure must have a unique colour gradient while retaining disparity between each other. A possible method for colour selection is discussed in detail in section 3.2.2.3.

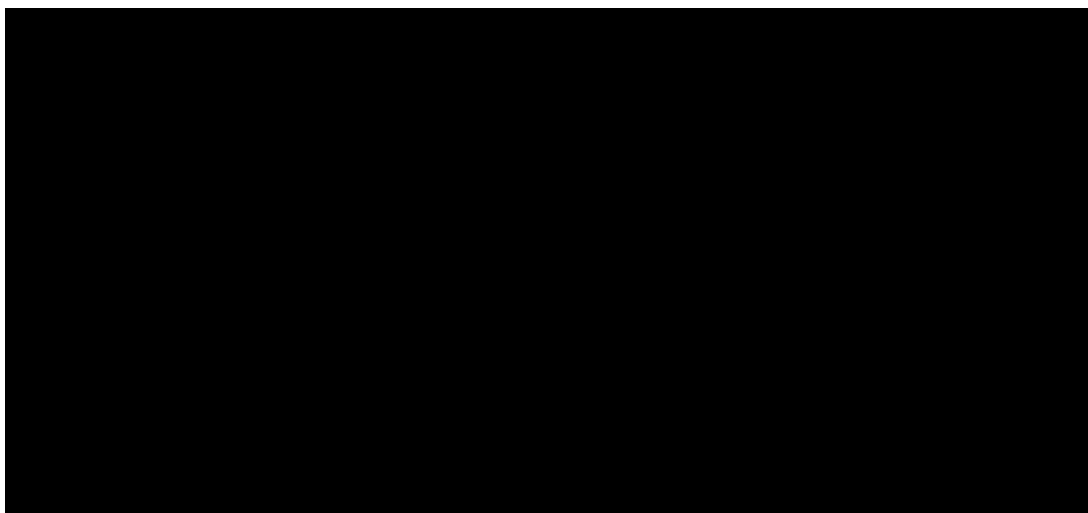


Figure 3.1: Basic Comparison of Camera View Modes.
Perspective View (Left) and Orthographic View (Right) [129].

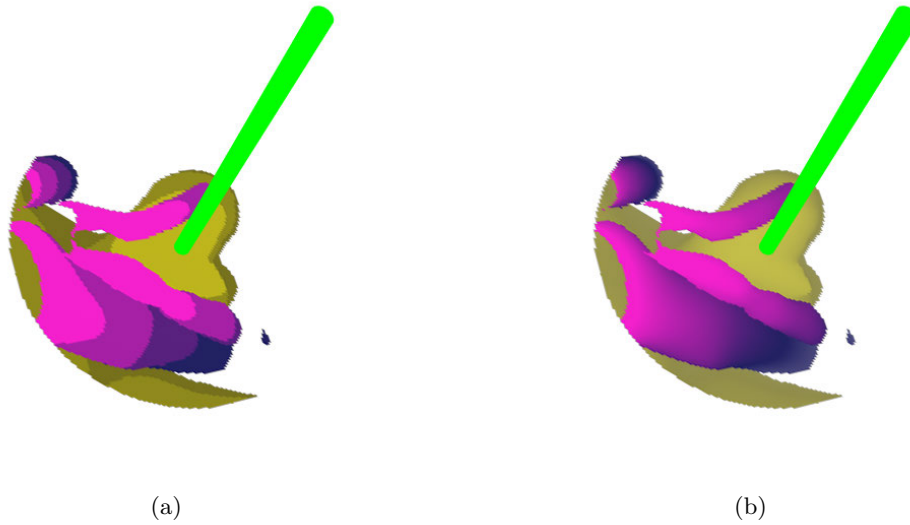


Figure 3.2: Two Gradient Methods.

a) Posterised, and b) Smooth. The depth perception might be hindered when the gradient is posterised.

In contrast to the virtual inspector, this method can exploit the orthographic mode's advantages. Since the depth information is now encoded in the colours of the objects, using the orthographic mode would expand the viewing zone. This expansion directly affects the speed of investigation, covering a larger area in a shorter amount of time. Additionally, due to the nature of the orthographic mode, the occlusion probability of the background objects by foreground content becomes minimal. Furthermore, the size of the objects is reported in the actual scale, which can relay a better understanding of the structure to the surgeon.

3.2.2.3 Colour Selection Procedure

As mentioned earlier in multiple instances, while designing an augmented reality system, one should always consider which information to overlay, and at what cost. In our specific case, we are trying to overlay a visualisation on a video stream that is coming directly from the inside of the patient's body. In this stream, one can detect many details such as landmarks, bleeding, instruments, implants, and abnormalities. In this situation, if the visualisation occludes any of this visible critical information, it can lead to severe complications further down the line. Moreover, given that our surgical environment has a constraint on colour space, one must be extremely mindful of choosing the appropriate colour. This colour should bring the situation to the user's attention while keeping a noticeable contrast from the environment's colours, is intuitive to interpret, and is blocking the least amount of (already existing) information in reality. As discussed earlier, the panel in method one and the colour map in method two are (adjustably) transparent which enables the user, even though not wholly, to detect information in video stream which lies behind the overlaid content. As for determining the correct colour set, an approach based on the outcome of related research in colour space and colour perception could be exploited.


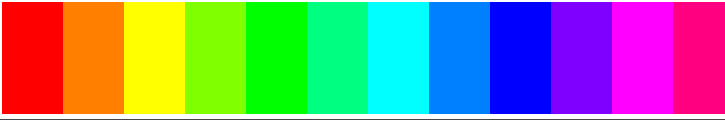
| | |
|----------------|--|
| Dark Colours |  |
| Bright Colours |  |

Table 3.4: List of the Colours Used as an Overlay.

When discussing the colour space, inevitably, an association with the various colour models that exist, e.g. RGB, CMYK, HSV, and HSL, is formed. Each of these models has roots in a single theory based on human colour perception [130]. In recent years with the rise of computer vision, Hue-Saturation-Value (HSV) and Hue-Saturation-Lightness (HSL) have shown great promise in closing the gap between human perception and computer vision. These models have proven to imitate human colour recognition and perception with better accuracy than the previous models when used in machine vision [131]. This outcome is due to the ability of these models in explaining how the human visual system gives more weight to the value and saturation while trying to distinguish between colours [131–133]. This finding can imply that, theoretically, if we analyse an image in HSV colour space, we can recreate the human recognition and differentiation of that colour.

A standard approach to analysing an image in computer vision is to extract its histograms and then investigate the extrema and changes based on the histogram values [134, 135]. In turn, to analyse the difference between multiple images, one can analyse their histogram distance. Common methods that are used to calculate histogram distances are Bin-to-bin, e.g. Chi-square Distance and Bhattacharyya, and Cross-bin methods, e.g. Earth Mover’s Distance and Mahalanobis [134, 135]. In this case, given the results from a study by Marin-Reyes et al., Chi-square distance is implemented [135].

Based on these arguments, two images from two different laparoscopic surgeries, with different instruments and conditions (no bleeding, bleeding), were taken (fig. 3.3a and 3.3d). These images then were overlaid 24 times with various transparent colour circles (e.g. fig. 3.3b, 3.3c, 3.3e, and 3.3f). These colours can be generally described as 12 dark colours, to be used as the minimum in the colour gradient, and their (12) bright counterparts, as the maximum of the colour gradient (see table 3.4). In the end, 12 variations of each image with dark colours overlay (total of 24) and 12 variations of each image with bright colours overlay (total of 24) was produced, leaving a total of 50 images (including two with no overlay).

The images were then imported to Matlab [136] and converted to the HSV colour space with each channel (hue, saturation, and value) being separated. Next, for each image, a histogram of each channel was calculated. Afterwards, the histogram distance between the original image and the overlaid image was computed for each channel (fig. 3.4). The differences have been plotted in a bar graph heralding the highest contrasting colours in the dark (fig. 3.5) set. Based on this,

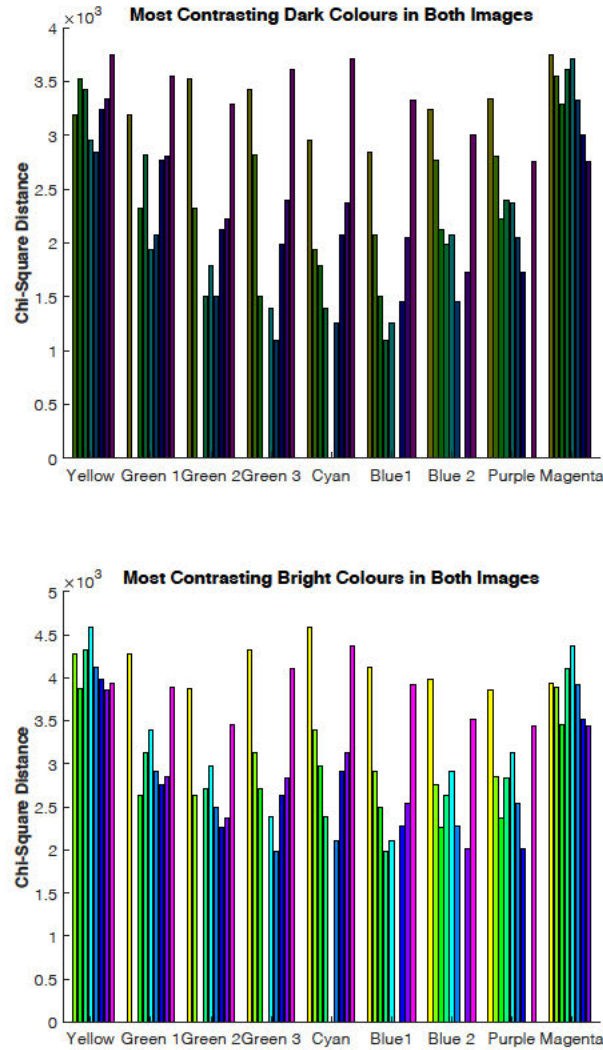


Figure 3.5: Chi-Square Distance in between the Overlaid Colours.

The averaged distance between two images for each colour is depicted for dark colours (upper), and bright colours (lower).

Hence, given the fact that all the users of this solution are surgeons and medical students, the established background of the colours in medical textbooks can be exploited. This method ensures that the user will have minimum difficulty determining the colour representation of our structures. While veins are mostly depicted as blue in the textbooks, the red vascular system can pose complication in our presentation. Hence, an alternative colour that is similar to red should be selected. The urinary system is mainly shown in yellow colours.

Accordingly, for the Pseudo-chromadepth mapping, dark and bright yellow colours were chosen to represent the renal collecting system's gradient. Dark blue and bright magenta were selected for the renal vascular system. In the virtual inspector, the colour yellow for the collecting system, and the colour blue for the vascular system were selected. The ability to change these colours, to other possible colours (in the colour space), at any point is also provided.

3.2.3 Auditory Display

As discussed earlier in section 2.3.2, the AD delivers the information to the user’s auditory domain. One of the advantages of this multidimensional information stream is that it can reduce the user’s uncertainty and provide a sense of confirmation. Since we are aiming to deliver a set of information containing multiple parameters, the parameter mapping seems like a sensible approach.

To address the core features of the design (section 3.2.1), the presentation of the data in a continuous and non-discrete type could challenge the usability of the system. For instance, if the spatial information were to be provided in a continuous format, e.g. continuous changes in the pitch, the estimation of these values by the user requires constant comparison. The user needs to oscillate around the target location to comprehend the relativity of the information, and this could affect the duration of the task completion. Hence, to make the auditory feedback more comprehensible and intuitive to interpret, a categorisation is necessary. An overview of these categories is available in table 3.5.

While the type of structure is already separated into two classes, i.e. renal vessels and urinary tubes, spatial information still requires classification. To this end, we first determine the spatial information which should be mapped in our auditory system. Two sets of information stand out the most among other candidates, namely the size of the structure and the distance from the instrument’s tip.

In contrast to VD, the identification of size is not an instinctive treatment in the auditory

| Type | |
|-------------------------|--|
| Renal Vessels | Including the renal artery and vein outside, and all the segmented vessels inside the kidney. |
| Renal Collecting System | Including Ureter, renal calyces, and medulla. |
| Distance | |
| Inside | Distance to the structure is less than $1mm$. |
| Close | Distance is between 1 to $5mm$. |
| Far | Distance is between 5 to $15mm$. |
| Out of Sight | Distance is more than $15mm$. |
| Density | |
| Low | Structures such as a single vessel branch, interlobular arteries, papilla, and minor calyces. |
| Medium | Vessel branching points and edges of the medulla. |
| High | Renal artery branching point, locations such as the renal pelvis with a high concentration of vessels and urinary tubes. |

Table 3.5: Discrete Auditory Parameters Categories.

domain. Hence, a clear definition of the size must be established, so that the users can all receive universal feedback regarding this variable. In our setup, in addition to dimensions, the complexity and density of the structure are remarked as size. For the sake of terminology, the size variable will be referred to as density from now on. To elaborate, a branching point in a vessel tree, the number of vessels in the scanning zone, and the diameter of the vessel impact the density variable. Three classes are identified for this variable, i.e. *small*, *medium*, and *large*. These classes offer more variability and granularity for our feedback while considering the processing speed of the information. To get a better understanding of these categories, one can remark interlobular arteries as small structures, while renal artery and medullas are considered as large.

The instances of using auditory mapping to report distance, in various tasks, have been mentioned in section 2.4.2.2. This parameter has been established as a crucial factor in many PM auditory systems, ours included. The usage of discrete values could benefit this feedback as well. To this end, similar to the density, four categories are defined, namely *out of sight*, *far*, *close*, and *inside*. It should be stated that for the first class, i.e. out of sight, no auditory tone is generated; hence the silence represents this class and the absence of sub-surface structures. Regarding the other three classes, a distance between 5 to 15mm is considered *far*, range of 1 to 5mm is *close*, while less than 1mm would be flagged as *inside*.

Based on the anatomy of the kidney in reality and our model, the presence of two types of substructures in the same area is inevitable. Hence it is of utmost importance, while designing each method, to ensure that the chosen tones should be easily distinguishable, specifically in scenarios with two types of structures in front of the instrument tip. In all the methods, the frontal structure would be heard unfiltered, where the posterior structure would receive a low-pass filter, replicating the natural effect of the sound when being obstructed. The tone that would be heard from the rearward structures contain information about their density (and not their distance). It should be stated that the low-pass filter should not affect the acoustic property of the density, that might lead to misclassification (e.g. the pitch).

Under the consequences of the high level of noise in the operating room [120], i.e. approximately 65dB in laparoscopy and nephrectomy [121], the sonification should ensure that the implemented tones are intuitive and easily differentiated from the background sounds while keeping unique characteristics that would separate them from the already existing auditory cues, e.g. alarms.

Another design factor to be considered is the spatialisation of the sound. Whether the sound is provided in monophonic or stereophonic environments, can impact the perception of the user. The monophonic audio generation is simple to implement and can be delivered through the majority of the audio output devices. In this method, the audio sources are generated and heard from one position (i.e. middle). However, in the stereophonic presentation, the spatial location of the generated tone can be manipulated. This method provides a better understanding of the spatial relations of the audio sources. A compatible output device is required to deliver this spatialisation successfully. Examples of these devices are stereo speakers and headphones. While our system can benefit from a stereo design, ensuring the availability and practicality of

such devices in an operating room, can be challenging. Hence, all the AD designs are based on monophonic audio reproduction.

Finally, one should bear in mind that people have different hearing capabilities and that each person might not be able to hear every tone wholly, or at least similarly. Many auditory ranges have been suggested and implemented in audio devices, aiming to reduce the effects of this issue. One of the more common of these ranges is Voiceband or Narrowband that is used in telephone communications. The Voiceband limits the frequencies in between $300Hz$ and $3.4kHz$. An alternative bandwidth that has been suggested more recently is Wideband, where the frequencies between $50Hz$ to $7kHz$ is transmitted. Having an auditory tone that remains within, at least, one of these ranges, could significantly increase the chances of it being successfully heard.

Eventually, a total of three methods (four designs) were conceptualised, which will be discussed in the following subchapters.

3.2.3.1 Method One – Auditory Icon and Synthesised Tone

With the underlying factor of highly contrasting tones in mind, two sets of sounds are to be selected to represent the types of the structures. A possible approach to a more efficient auditory recognition is the utilisation of auditory icons. Appointing an auditory icon as one of our tones could ensure the recognition and distinction of the tone from the other auditory sources. This differentiation can be boosted furthermore by implementing an absolute polar approach to an auditory icon, a synthesised sound. This high contrast should ascertain a more rapid distinction between the types, and later on, a more feasible concentration on each of the sound sources.

To locate the most associable auditory icon in the given scenario, one can reflect on daily life experiences. The sound of running water can extensively replicate the sound of the liquids existing in our renal system, i.e. blood and urine. However, the association of water sound and urine is more intuitive and commonplace in daily life (even to an unconscious level) than that of water and blood. Hence, the sound of running water is selected to represent the collecting system. Even though this selection might be considered humorous in some cases, this can further improve the associability of the sound and the structure.

As for the counterpart to the collecting system, a computer synthesised tone should be created to represent the vessels. The primary assumption that should be adopted in the design of this tone is that it must have a unique and distinctive characteristic compared to the sounds that are present in the operating room, e.g. alarms.

Since the PM is the preferred method of sonification, the following acoustic characteristics are considered as representatives of our variables. In the synthesised tone, the pitch indicates the density of the vessel, with higher density being reported with a higher pitch. Furthermore, the distance is mapped to the rhythm. Similar to the Geiger–Müller counter, a faster rhythm represents a lesser distance. While reporting a *far* structure, the tone remains continuous and without any cut-off.

Due to the absence of a (distinct) pitch in the water sound, the flow pressure is manipulated to report the density of the urinary structures. With a higher density of urinary structure, higher pressure in the flow is heard. To be able to induce a rhythm in water flow, a cut-off in the continuity of this flow is proposed. This discerning factor can be further improved by recreating the sound of water splashes or waves robbing against the shore. The speed at which the waves are repeated can be an indicator of rhythm, and in turn, can report the distance variable. Similar to our synthesised tone, a higher rhythm depicts a lesser distance.

3.2.3.2 Method Two – Musical Instruments with a Broad Tempo

In an alternative method, instead of an auditory icon and a synthesised sound, musical instruments are utilised. The definitive and comforting nature of the musical instruments could improve the user experience with the setup; hence, two musical instruments are to be selected to represent each of our structure types. Both instruments benefit from the clearly defined pitches of musical notes.

Timbre is one of the main reasons behind the human ability to distinguish between sound sources and musical instruments [60]. Furthermore, the creation method for each sound contributes heavily to its characteristics and timbre [60]. In other words, a plucked string instrument is very distinguishable from a wind instrument or a bowed string instrument. According to this argument, a plucked instrument, e.g. guitar, a bowed instrument, e.g. violin, a wind instrument, e.g. flute, or a percussion instrument, e.g. a drum set could all be possible candidates. However, percussion instruments might not be easily differentiated from the present sounds in the operating room. These instruments can also be distracting. Hence, to avoid this outcome and maintain a uniform instruments composition (i.e. orchestral instrumentation), harp, viola, and flute are selected as the primary candidates.

Similar to the previous methods, the structural density is reported through pitch fluctuations, while the distance is mapped to the rhythm of the tones. However, the main difference in this method lies within the repetition speed of the tones. In this concept, each note is played for a short period while followed immediately by a silent period, with the same duration. The underlying objective of this design is to provide a buffer time for the user during which (s)he could process the received information without being affected by the sensory or mental overload. During this buffer period, the auditory channel remains semi-open, offering a chance to maintain contact with the real world. Given the constant communication of the surgeon and their assistants during the surgery, this buffer time can prove specifically advantageous.

3.2.3.3 Method Three – Music Generation Cloud

The final AD method is a music generation cloud which constantly creates a randomly generated series of tones as if listening to a musical piece. In this method, similar to the former approach, each of the two musical instruments represents a particular structure type. Similar to a concert, the two instruments play a random melodic progression of notes while keeping harmonic, yet distinct, sound characteristics.

As a consistent approach among all AD methods, the density is mapped to the pitch, and the distance is mapped to the rhythm. The random melody is generated from a series of notes in a same musical octave, suggesting that instead of having the musical note's pitch representing the density, the pitch of the musical octave represents this variable. As the density of the structure increases, the octave of the tone is shifted towards higher pitch notes.

As mentioned earlier, this system resembles a musical piece; hence rhythm of the piece is the speed at which the musical notes are being played. The faster the notes are being played, the closer the instrument tip is to the target and vice versa. In case that two structure types are present under the instrument tip, the rear structure's instrument is heard less dominantly, enabling the frontal instrument to act as a solo instrument in a classical concert.

The goal that this design follows is the presentation of a more aesthetic and pleasing sound that the user can listen to for a more extended period without getting frustrated. The random and non-repeating nature of the generated melody could defuse the irritation experienced by the user. In parallel to providing the information, this relaxing tone can also act as a workload reducing factor, as suggested by [137].

3.3 Implementation

3.3.1 Development Environments

The development of the final intractable version of this dissertation was implemented through a unison of multiple hardware and development environments. An overview of the general software architecture of the solution can be found in figure 3.6.

Inclusion of NDI Polaris Spectra optical tracking system [138] permits a real-time interaction between the software and the laparoscopic instruments. This system uses infrared pulses to track passive fiducial markers attached to our instruments. In this environment, three objects were tracked,

- the world anchor, representing the static point of our tracking system,
- the laparoscopic camera, and
- the interaction instrument, in this case, a grasping forceps.

The world anchor is a 3D printed model of the kidney with which the user can interact. Using fiducial markers, this model is tracked and is considered as the static point of our tracking system. Every other movement tracked by the system is relative to this fixed point. The kidney model is fixed to the table to prevent any movement and dislocation. The generated content is registered and overlaid on the tracked objects with respect to the information reported by the tracking system. The rendered volume and surfaces of the kidney's structures are placed virtually at the position of the 3D printed kidney model. To be able to acquire the laparoscopic video stream, the B. Braun EinsteinVision® 3.0 [139] laparoscopic system was used in the setup of this project.

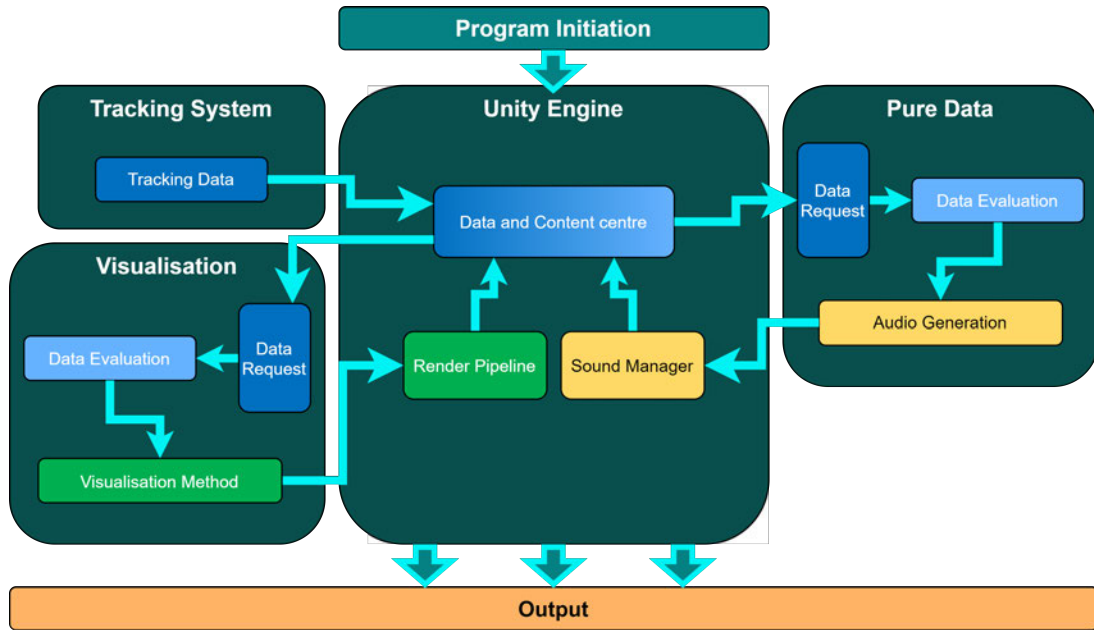


Figure 3.6: General Software Architecture of the Solution.

Unity Engine is the central hub for handling the information. NDI tracking system supplies the tracking information to Unity Engine. Unity Engine then transfers this data to our visualisation and sonification modules. The sonification is done via Pure Data, while visualisation is realised with Unity Engine.

The development of AD was conducted through Pure Data (Pd) [140], a visual programming language dedicated to creating interactive real-time music systems to be used in a multimedia platform. Pd utilises a series of modules and objects in a simple GUI to help the users create their desired sound systems (fig. 3.7).

The VD of this work was implemented mainly through the Unity Engine (version 2018.3.0f2)

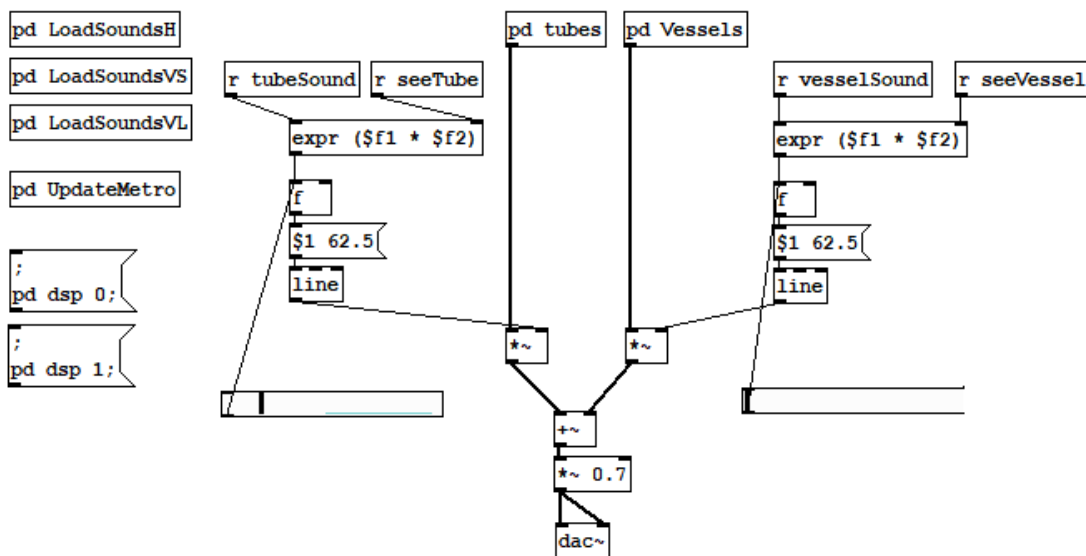


Figure 3.7: An Example of Pure Data (Pd) Environment.

[141] with a personal license. The Unity Engine also served as the central integration environment where Pd and tracking system component are incorporated so that this prototype can be realised. The integration of the tracking system and unity engine was already available as a part of the working setup in the laboratory. To be able to have a functional pure data environment within unity, an open-source wrapper named LibPdIntegration was used [142].

To create a virtual 3D model of our organ, in our case a kidney, a segmentation of a healthy left kidney on an anonymous set of medical images [143] was realised via MeVisLab [144]. The segmentation was then converted into an STL model using 3DSlicer [145,146]. To acquire a more uniform and smoother model of the kidney, a series of smoothing and mesh repairs were carried out via Blender [147].

The specification of the computer, which the prototype was built and tested on, can be found in the following table (3.6),

| | |
|------------------|--|
| Processor | Intel® Core™ i7-4790K CPU at 4.00GHz |
| Installed RAM | 16GB |
| Graphic Card | NVIDIA® Quadro K2000 |
| Graphics Memory | 2GB |
| Operating System | Microsoft® Windows 10 professional 64-bit. |

Table 3.6: Specifications of the Setup Computer.

3.3.2 Visual Displays

Since the only required information lies in front of the instrument, all the content that is present behind the instrument tip is unnecessary. A clipping plane is placed perpendicular to the instrument tip, omitting the redundant information from the VD.

3.3.2.1 Virtual Inspector

In this VD, a virtual camera was placed at the tip of the instrument, viewing our recreated 3D model of the organ and recasting the image on to a small movable screen within our viewport (fig. 3.8). Given the already visualised model of the kidney, necessary information, in respect to the task, can be located on this visualisation, e.g. the vessels and urinary tubes.

As the instrument traverses the area, one can witness the structures that precisely lie beneath, i.e. those which are directly in front of the virtual camera. The camera sees only up to a certain adjustable distance. This distance can be decided by the surgeon, depending on the suturing task. This constraint is also applied to the viewport's width and height, in which the camera has a limited field of view, narrowing down the information that the user receives. The reason behind these constraints is to eliminate excess information in the viewport of the operation to keep the mental load of the user and occupation of the screen space minimal. A green ring around the instrument tip designates the outline of the viewing zone.

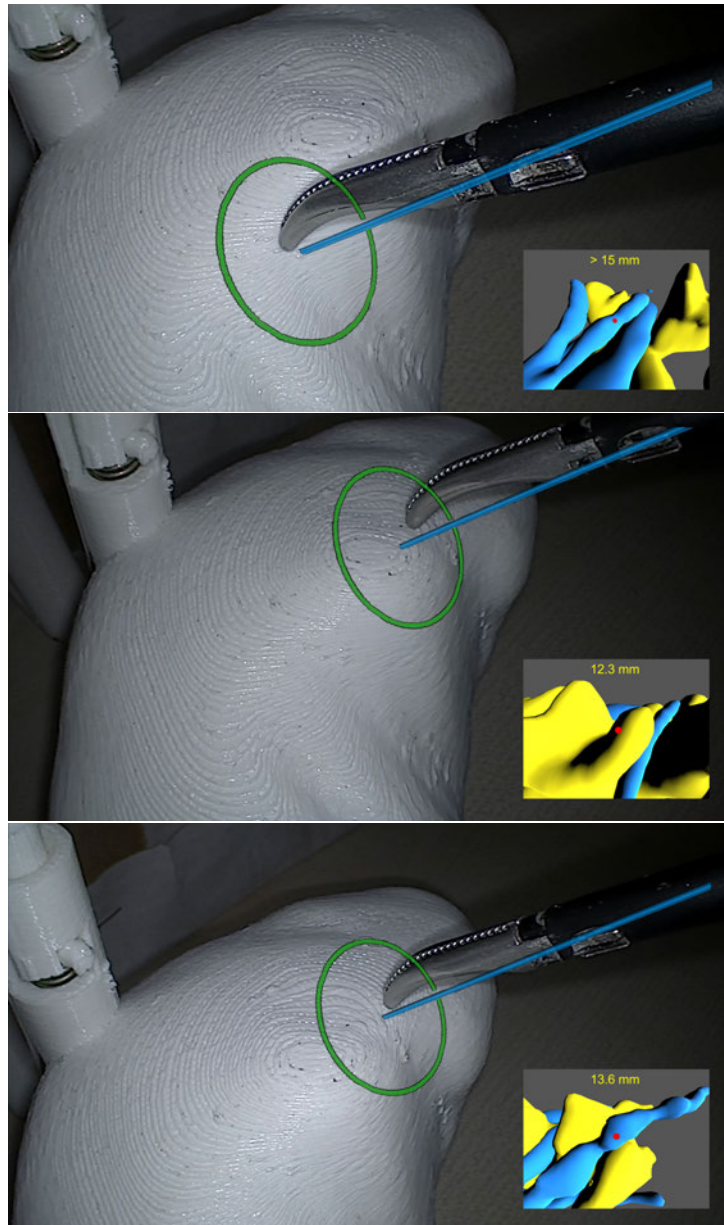


Figure 3.8: Virtual Inspector in Use.

The renal vascular system is shown in blue, while the collecting system is shaded yellow. The closest point on the mesh is shown with a red dot, and the distance to this point is reported in the textbox (e.g. 13.6mm).

In addition to spatial information acquired directly from the image, a textbox, which is located on top of the panel, continuously reports the exact distance between the virtual instrument tip and the closest point on the facing surface of the kidney mesh. The textbox has, in total, four different scenarios,

- Structure more than 15mm away reported with " $> 15\text{mm}$ ",
- Structure $d\text{ mm}$ away with $1 < d < 15$, reported with " $d\text{ mm}$ ",
- Instrument tip is inside a vessel, reported with *Inside Vessel*, and

- Instrument tip is inside a urinary tube, reported with *Inside Tube*.

A sphere is cast forward from the tip of the virtual instrument in each frame, announcing the collision whenever contact is made with a vessel or tube within our model. The traversed distance of the sphere up to the contact point is stated as the distance between the structure and instrument tip. Given the fact that we already know our computer-generated volumes and surfaces, the closest point on the structure from the instrument tip can be acquired through the same method. The first point, with which the sphere comes in contact, is marked as the closest point. This sphere only penetrates a user-defined maximum value. This limit could abridge the investigation process by reporting only the information that the surgeon finds necessary. If the suturing is to be applied with a depth of 1cm from the surface, reporting everything that lies beyond 3cm point inside the kidney can possibly distract and confuse the surgeon, with potentially no positive outcome.

Implementing a sphere cast, instead of a ray cast, benefits us with the possibility of scanning a larger area in front of the camera. This approach also takes into account some of the possible errors that might occur in the setup, e.g. registration, segmentation, and surface model creation errors, and unwanted movements/jitters that might occur because of the tracking system and the user's movement. An alternative to the sphere cast is the cube/plane cast. This method could prove to be as effective as sphere casting; however, there are few issues regarding this technique. The most dominant shortcoming of cube casting is that it fails to perceive information on objects that are on a curve. Given that our structures of interest are tubular or curvature, the initial point that is touched by a cube would be the front face (relative to the instrument tip), while the closest point might be lying, on the curve of the vessel, lower than its front face. Second, the cube cast does not report equidistant points in front of our tip; but depending on the position of the point, and collision with the cube cast, this distant varies.

A small red sphere instances a laser-pointer, which reports the position of the closest point found on the structure (from which we can see the distance). Furthermore, the penetration depth of the sphere cast can be manipulated at any time before or during the runtime.

3.3.2.2 Real-time Pseudo-chromadepth Mapping

This method generates a real-time Pseudo-chromadepth image at the tip of the virtual instrument, which helps the user with recognition and localisation of the renal structures. The distance from the instrument is mapped onto a colour gradient, which represents the closer with brighter and further away with darker colours. Additionally, a green ring around the instrument tip indicates the boundaries of the colour map.

The principle behind orthographic imaging is the planar projection of the scene. In other words, rather than looking through a pinhole, the scene is viewed through a series of equidistant planar points. Each of these points can cast only one ray in the forward direction. To be able to reproduce this effect in this design, instead of investigating the distance from one point, e.g. the

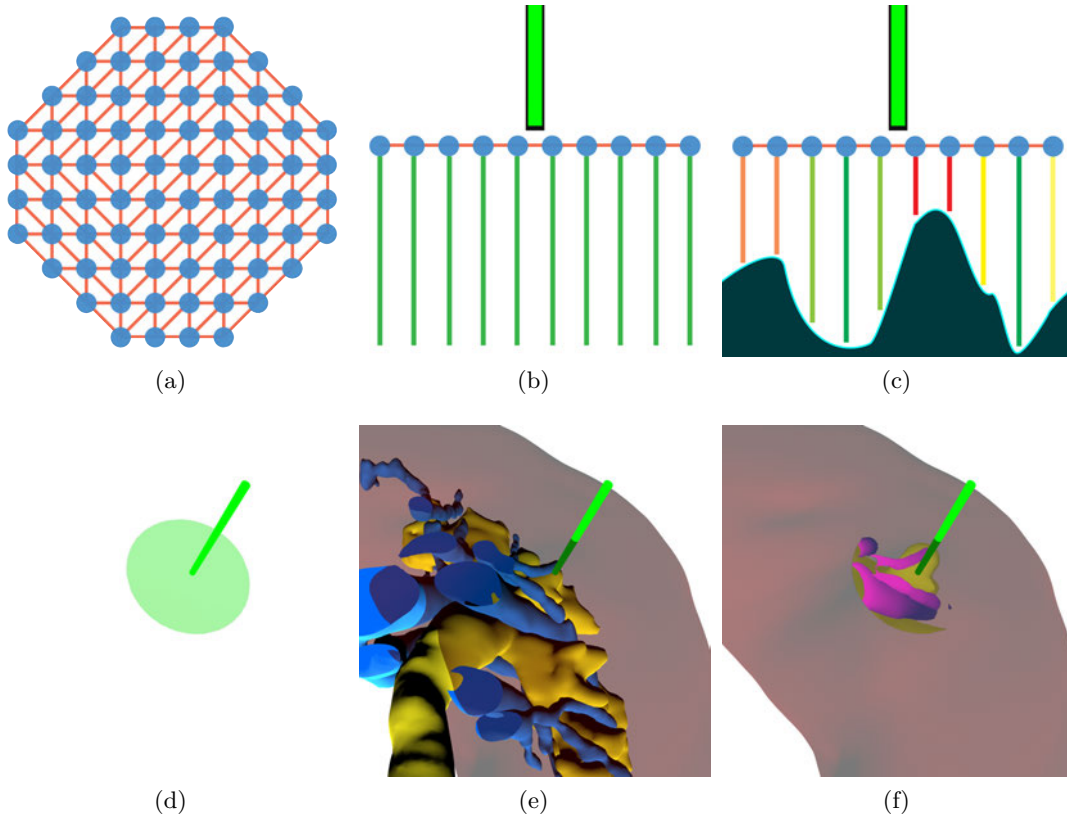


Figure 3.9: Pseudo-chromadepth Mapping Principle and Implementation.

The concept of a) the circular Pseudo-chromadepth mapping mesh, b) the ray casting of each vertex, and c) the colour-coding of the distance is depicted in this figure. The implemented version of d) the circular mesh, e) the renal substructures that are to be colour-coded, and f) Pseudo-chromadepth map are shown as well.

instrument tip, the distance of the objects from a plane centred at an ideal point (e.g. instrument tip) should be measured.

Accordingly, the camera was replaced with a generated mesh placed perpendicular to the instrument shaft's axis (with the instrument tip being at its centre). This mesh is formed by an $n \times n$ net of vertices, which ultimately shapes a plane mesh. To reduce the screen space occupation, the mesh is reduced to a circular form (fig. 3.9a), surrounding the instrument tip. Each one of these vertices casts a ray forward (according to instrument's orientation), at the runtime (fig. 3.9b). The distance that each ray traversed from its vertex to the structure's surface, which it collided with, is then returned and normalised between the preset minimum and maximum depth investigation value. The normalised distance is then interpolated into a value from zero to one on a colour gradient (fig. 3.9c), creating a colour-coded orthographic projection of what lies under the tip. This projection is displayed on the generated mesh surrounding the instrument tip. The colour gradient starts emphasising the presence of structure by becoming less transparent as the objects draw near. Furthermore, the brightness and saturation of the colours increases as the distance between the objects and the generated mesh becomes less. This effect is known as aerial perspective, or fog as seen in the work of *Kersten-Oertel et al.* [103].

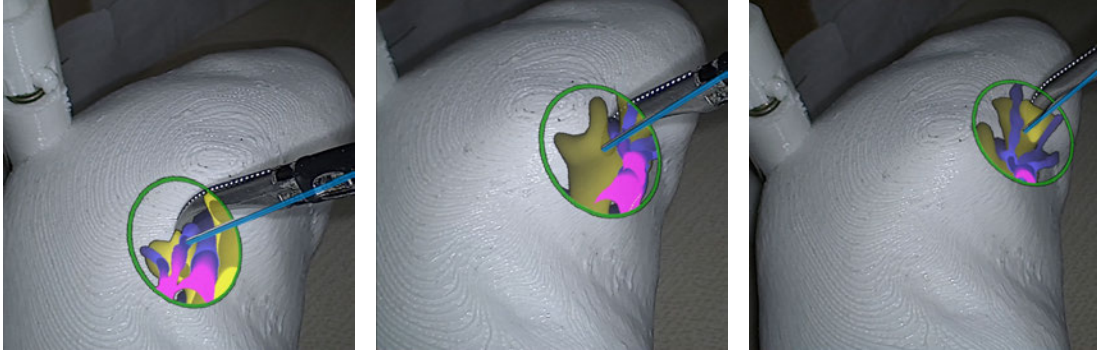


Figure 3.10: Real-time Pseudo-chromadepth Mapping in Use.

The urinary collecting system is colour-coded with a gradient of yellow, while the vascular system is coloured with a gradient from blue to magenta.

It is fairly intuitive that with the presence of two structure types, a separate colour gradient should represent each structure. For this purpose, based on the structure that each ray collides with, whether the vessel or the urinary tubes, the reported distance is evaluated via a separate colour gradient. The result of this implementation can be found in figure 3.10.

Similar to all the methods, depending on the dimensions and position of the resected site, the investigation depth can be adjusted. The range, in which the colour map should report the minimum distance to alert the user, can also be adjusted to suit each case specifically. As for the colour map, the size, number of vertices, its shape, and gradient colours can be adjusted to the user's preference. The adjustable options for this method are provided in the appendix (fig. A.1).

3.3.3 Auditory Displays

In every auditory method, the density is processed via counting the number of vertices that are reporting a collision with a specific structure type, similar to the Pseudo-chromadepth method but in a smaller radius. The distance from the target is calculated with the help of sphere casting, similar to the Virtual Inspector.

All the sounds and tones used in the following methods were created, recorded, and processed by the author of this thesis. These sounds consist of one alert tone, one interactive tone generator, six water variants, 15 Harp, 30 viola, and 30 flute tones, summing up to a total of 83 tones.

3.3.3.1 Auditory Icon and Synthesised Tone

In this method, two highly contrasting sounds were selected to represent the structures. First, the sound of running water was selected to represent the collecting system, and on the other hand, a computer synthesised tone was created to represent the vessels. The density of the structure is mapped to the water pressure for the collecting system and the tone's pitch for the vessels, with higher pressure and pitch indicating a denser structure. Finally, the rhythm of each tone is a translation of the distance between the instrument tip and the closest point on the targeted structure, with a faster rhythm representing lesser distance.

To express the density of the collecting system, the water pressure is manipulated to produce three conditions, i.e. low, medium, and high pressure; representing *low*, *medium*, *high* density. Since this metaphor might prove counter-intuitive to some users, the reversed mapping of it is also available.

The water tone is triggered every 250, 500, and 2000ms, depending on the distance, i.e. *inside*, *close*, and *far*, respectively. The amplitude of the water sound is modulated to create a water wave sound effect. The modulation multiplier ranges from 0.3 to 1. With respect to the period of the trigger, the duration of this modulation differs. This modulation can be explained in the following formula,

$$A = m(t) \times A_w, \quad (3.1)$$

$$m(t) = 0.35(\cos(2\pi Tt)) + 0.65, \quad (3.2)$$

Where A is the amplitude of the output, A_w is the amplitude of the water sound before modulation, $m(t)$ is a time-dependent multiplier of the modulation, and T is the trigger frequency per a second. For instance, for the trigger repetition of 250ms the T value is 4.

The closer the structure gets to the instrument tip, the faster the waves would become. More specifically, a *far* structure resembles an uninterrupted flow of water, a *close* structure is depicted as rhythmic splashes in the stream, and being *inside* the structure has more rapid splashing rhythm accompanied by an alert sound with a frequency of 739.989Hz (F#5 musical note).

On the other hand, the synthesised sound embodies the maximum contrast to the water sound by being a digitally generated tone, (theoretically) bringing an assured distinction between two sounds. The synthesised sound is created from the base frequencies of 65.4Hz, 130.8Hz, and 261.6Hz (C2, C3, and C4 notes), and harmonised by each frequency's first to eighth harmonics, creating a complex tone. A complex tone is when a pure tone (single frequency sound wave) is accompanied by its harmonics (overtones), leading to the formation of a unique character for the sound (depending on the overtones). The frequency of each harmonic can be easily computed using the given formula,

$$f_h = h \times f_0, \quad (3.3)$$

With f_h being the frequency of the harmonic, h is an integer representing the number of the harmonic, and f_0 is the base frequency. Since the human perception of loudness is not a constant nor a linear function, there is an exponential decay designed into the harmonic frequencies' loudness, decreasing as the frequencies go higher, which characterises the sound as well as maintaining the emphasis on the base frequency [60].

In addition, each complex tone is complemented with 7th, 12th, 16th, and 19th semitone intervals, adding more character to the created tone and forming a musical chord. To calculate

the frequency of each interval, one can use the following formula,

$$f_n = f_0 \times a^n, \quad (3.4)$$

where f_n is the frequency of the n^{th} semitone interval, f_0 is the base frequency, a is $2^{\frac{1}{12}}$, and n is the number of the semitone interval.

The density of the vessels is determined by the pitch of the tone, meaning that 65.4Hz , 130.8Hz , and 261.6Hz (C2, C3, and C4 notes) represent *low*, *medium*, and *high* density, respectively. The number of times that each tone is played expresses the distance between the instrument tip and the closest point on the targeted vessel. Similar to the water sound, a continuous tone represents a *far* vessel, while a *close* vessel is heard as the tone being repeated every 500ms with a duration of 400ms . Eventually, being inside the vessel triggers the alert sound, playing every 125ms accompanied by the tone every 250ms with the duration of 200ms (identical to the water sound paradigm).

3.3.3.2 Musical Instruments with a Broad Tempo

As discussed in section 3.2.3.2, two musical instruments were selected so that each represents a structure, i.e. harp for the collecting system and viola for the vessels. The density of the structure is mapped to the pitch, while the distance is mapped to the rhythm of each musical note. The mapping paradigms are identical to the synthesised tone from the previous method.

The harp tones have the base frequencies of 65.406Hz , 146.832Hz , and 466.164Hz (C2, D3, B_b4 musical notes) from the low to the high pitch, and density, respectively. The Viola tones follow the same base frequencies.

The chief aspect of this design is the duration of the tones. For a structure that lies in the *far* category, the tone is played once every 2000ms for a duration of 250ms . For *close* structures, the tone is heard once per 1000ms and for a duration of 250ms . Finally, when *inside* a structure, the tone would be heard once every 500ms for the duration of 250ms , accompanied by the alert tone, similar to the previous method.

3.3.3.3 Music Generation Cloud

The final AD method is a music generation cloud. In this method, similar to before, two musical instruments were selected. The harp was selected to represent the collecting system, and the viola was selected to express the vessels. The density and distance mappings are similar to previous methods. Alternatively, the flute was used as a substitute to viola, representing the vessels.

The tones for each density are melodies created from a rhythmic progression of four randomly selected based frequencies (or musical notes) from a pool of five (or a pentatonic scale) frequencies in a defined range (or a musical octave) and without repetition. As the structure's density increases, the octave of the note pool is shifted, with higher octave representing the higher density.

| Instrument | Density | Base Frequencies from Low to High in <i>Hz</i> | | | | |
|------------|---------|--|---------|---------|-----|---------|
| Harp | Low | 65.406 | 73.416 | 87.307 | 110 | 116.541 |
| | Medium | 130.813 | 146.832 | 174.614 | 220 | 233.082 |
| | High | 261.626 | 293.665 | 349.228 | 440 | 466.164 |
| Viola | Low | 65.406 | 73.416 | 87.307 | 110 | 116.541 |
| | Medium | 130.813 | 146.832 | 174.614 | 220 | 233.082 |
| | High | 261.626 | 293.665 | 349.228 | 440 | 466.164 |
| Flute | Low | 130.813 | 146.832 | 174.614 | 220 | 233.082 |
| | Medium | 261.626 | 293.665 | 349.228 | 440 | 466.164 |
| | High | 523.251 | 587.33 | 698.456 | 880 | 932.328 |

Table 3.7: Music Generation Cloud Base Frequencies.

Each base frequencies row represents the musical octave used for representing the mentioned density.

As the instrument tip gets closer to the structure, the speed at which these notes are played increases. This progression leads to a seemingly unending musical piece that is solely depending on the user's movement and structures' properties. The base frequencies of each one of these tones can be found in table 3.7.

If a structure is *far*, the progression plays a note every $500ms$ for the same duration. Alternatively, a *close* structure is indicated with a note being played every $250ms$, and when the instrument is *inside* a structure, a note is played every $125ms$. In order to make the viola and flute melodies more aesthetic and pleasing to listen, for notes that are played every $500ms$, a more sustained style of instrument playing (also known as legato) is employed. On the other hand, for the notes with a trigger frequency of $250ms$ and $125ms$, the playing style is short and quick (also known as staccato).

3.4 Discussion and Verdict

In the end, a total of five display methods were conceptualised and implemented, i.e. two VDs and three ADs. After realising all the concepts, the aptitudes of each became more apparent, giving us a chance to compare and discuss their gains and shortcoming.

Being able to see through the cortex and directly view the internal structures of the kidney appears as the appropriate technique. In addition to this view, the visual inspector provides spatial relations of the structures and the precise distance from the instrument tip. The position of the smaller viewport can be set by the user to ensure a full visibility of the surgical site. However, by setting this preview at the edges of the main screen, the user should frequently switch their visual focus from the surgical site, which is often located at the centre, to the outer regions of the screen. This shift of focus can affect the mental flow of the surgeon, and more so during this critical and time-sensitive phase of the surgery. Alternatively, this viewport can be placed

in the middle of the screen, or even at the instrument tip, similar to the Pseudo-chromadepth method. Even so, due to the specific lighting effects of the generated visualisation, the view might not be as pleasing as expected.

The Pseudo-chromadepth map could be considered as an expansion of the virtual inspector. This statement is due to the fact that the method also reports spatial information, distance, and see through the kidney's cortex. However, the precise measure of the distance was omitted in this method. Though the distance textbox offers valuable information, a rapidly changing number could be a distracting factor. Instead, an adjustable value was presented where one could adjust the maximum and minimum thresholds of the distance for colour interpolation. Due to the nature of the Chromadepth, the structural forms and the depth information are conveniently visible. In addition to this extension, the field of view is benefiting from an orthographic approach. As mentioned earlier, an orthographic mode can improve the field of view, as well as providing a more accurate estimate of the proportions and spatial relations of the objects; hence, providing more detailed feedback on a larger area in a shorter amount of time in comparison to the virtual inspector. In light of these arguments, the Pseudo-chromadepth mapping was selected as the default VD method.

Understandably, acquiring and calculating all this information can be computationally quite intensive and expensive. In the Pseudo-chromadepth method, the expensiveness is directly correlated with the number of vertices; that is, with a higher number of vertices, a higher calculation load is exerted on the GPU and CPU. Since this method is implemented in a game engine, the ray (sphere) casting is mainly done via GPU (unless faced with GPU limitations in which CPU becomes the central processing unit). With this in mind, the current setup comprises 160×160 vertices. This choice was a compromise between the performance and presentation. Intuitively, with lesser number of vertices, a more pixelated and less smooth colour map would be produced, and to lessen this issue, the number of vertices was pushed to the most reasonable limit in respect to performance.

As for the AD, With the goal of conveying a set of information, consisting of the type, density, and distance from the structure, all display methods were based on the PM *without reference* approach. Each of these designed methods provided unique benefits that cannot be compared in a general manner. This fact is even more hindering when considering the unique hearing process of each person.

Since the benefits of these methods were pointed out earlier in section 3.2, here, the shortcomings of each method are mentioned briefly. Based on the subjective opinions of the users, the water auditory icon can be distracting, counter-intuitive or even challenging to interpret at certain points. On the other hand, the short cues from the broad tempo method can be missed and require additional waiting time to be heard again. Finally, due to the extensive musical nature of the music generation cloud, if the user does not have a musical background, the interpretation of the information can prove demanding. Furthermore, since the solution of this project would be utilised under extreme time constraints, a prolonged listening session is not of concern in the solution's design.

In order to find the most suitable auditory design, an informal study was conducted. This study was performed with 12 participants, seven musicals and five non-musicals, with different backgrounds. Each participant experienced all three auditory methods, by explicitly being watchful of the accuracy and the duration of engagement with the method. In the end, they were asked to rank these three methods according to "the most amount of practical information acquired in the least amount of time" criterion. Respectively, the best method received three points; the second method was given two points, while the last method only received one point. For the final analysis, the points of each method were summed across all users. This analysis resulted in,

- the auditory icon and synthesised tone, 31 points,
- the music generation cloud, 21 points, and
- the musical instruments with a broad tempo, 20 points.

Built on this outcome, method one, i.e. the auditory icon and synthesised tone, was chosen as the default auditory feedback method.

4 Evaluation

4.1 Concept and Design

In the conventional LPN/RAPN method, the surgeon estimates the location of the renal structures based on the external landmarks that were recognised, up to that point, during the surgery. This approach can increase the probability of errors and complications during this phase and post-operation. Intending to decrease these undesirable outcomes, as stated in the research question, this thesis aims to investigate the improving effects of visual and auditory AR displays on the accuracy, duration, and workload of RSRP in LPN/RAPN.

To find a possible answer to this question, forming an associated hypothesis is considered as the first step. One of the more common approaches in scientific studies is using null hypothesis significance testing (NHST), which requires establishing sets of paired null (H_0) and alternative (H_1) hypotheses based on the study goals. In our case, the loose equivalent null hypothesis to the research question would be,

H_0 : Visual and auditory AR displays do not significantly improve the accuracy, duration, and workload.

Where the alternative hypothesis can be broadly stated as,

H_1 : Visual and auditory AR displays significantly improve the accuracy, duration, and workload.

The next crucial step in the process of designing a study is to determine the dependant and independent variables. Clearly, the causes and conditions that are to be investigated are highlighted in the research question, namely the VD and AD. The variables that we are aiming to measure and analyse in the hypothesis are the accuracy, duration, and workload. Hence, these variables are affected by our cause/condition, and they are established as our dependant variables. In this regard, the VD and AD are our independent variables.

Eventually, with the determination of the dependant and independent variables, one can provide more accurate and tailored pairs of null and alternative hypotheses:

- Auditory Display Main Effects,
 - H_{Aa0} : Averaged over the two visual display levels, the auditory display does not significantly increase the accuracy of the procedure.

- H_{Aa1} : Averaged over the two visual display levels, the auditory display has a significant increasing effect on the accuracy of the procedure.
 - H_{At0} : Averaged over the two visual display levels, the auditory display does not significantly decrease the duration of the procedure.
 - H_{At1} : Averaged over the two visual display levels, the auditory display has a significant decreasing effect on the duration of the procedure.
 - H_{Aw0} : Averaged over the two visual display levels, the auditory display does not significantly decrease the workload of the procedure.
 - H_{Aw1} : Averaged over the two visual display levels, the auditory display significantly decreases the workload of the procedure.
- Visual Display Main Effects,
 - H_{Va0} : Averaged over the two auditory display levels, the visual display does not significantly increase the accuracy of the procedure.
 - H_{Va1} : Averaged over the two auditory display levels, the visual display has a significant increasing effect on the accuracy of the procedure.
 - H_{Vt0} : Averaged over the two auditory display levels, the visual display does not significantly decrease the duration of the procedure.
 - H_{Vt1} : Averaged over the two auditory display levels, the visual display has a significant decreasing effect on the duration of the procedure.
 - H_{Vw0} : Averaged over the two auditory display levels, the visual display does not significantly decrease the workload of the procedure.
 - H_{Vw1} : Averaged over the two auditory display levels, the visual display significantly decreases the workload of the procedure.

These hypotheses are the basis for the inspection of main effects caused by each display method. However, the effects of each display on the other one are phenomena essential and worthy of investigating. To this purpose, the interaction effects hypotheses can be declared as,

- Interaction Effects,
 - H_{AVa0} : With respect to population means, regarding the accuracy, auditory and visual displays do not significantly interact.
 - H_{AVa1} : With respect to population means, regarding the accuracy, auditory and visual displays do significantly interact.
 - H_{AVs0} : With respect to population means, regarding the duration, auditory and visual displays do not significantly interact.
 - H_{AVs1} : With respect to population means, regarding the duration, auditory and visual displays do significantly interact.

- H_{AVw0} : With respect to population means, regarding the workload, auditory and visual displays do not significantly interact.
- H_{AVw1} : With respect to population means, regarding the workload, auditory and visual displays do significantly interact.

As stated in NHST approach, if the null hypothesis can be rejected with significant statistical evidence, one can assume that the alternative hypothesis holds. This process requires a valid and exact evaluation method that could thoroughly examine our hypotheses.

In order to design such an effective evaluation method, one must highlight the main structural components of the solution’s concept, i.e. the VD and AD. Each one of these components can provide sufficient information to aid the user in achieving the solution’s goal. Considering this, it is possible to breakdown the conventional and proposed solutions into four assistance categories,

- assistance without augmented reality, NoA,
- assistance with visual augmentation, VA,
- assistance with auditory augmentation, AA, and
- assistance with both visual and auditory augmentation, VAA.

Accordingly, it can be argued that each independent variable has two states,

- absent (Off), and
- present (On),

which results in two-level independent variables. The four assistance categories could be easily redefined using these independent variables, as viewed in table 4.1.

| | | Visual Display | |
|------------------|-----|----------------|-----|
| | | Off | On |
| Auditory Display | Off | NoA | VA |
| | On | AA | VAA |

Table 4.1: Assistance Methods and AR Displays State.

As discussed earlier, the conventional method used in LPN/RAPN resembles the first category, i.e. the absence of augmented reality. Hence, the NoA category can be used as a simulation of the conventional method (the control condition), while the other three methods can be used as alternative scenarios, representing the solutions offered by this thesis. In these scenarios, the user is aided in the localisation task with different AR displays.

Redefining the prementioned categories based on the independent variables provides the opportunity to establish a two-by-two study design. One significant benefit of a two-by-two

design is the possibility of investigating the effects for each augmentation approach, as well as their interactions.

Since this thesis aims to help the user with the recognition and localisation of the critical structures within the kidney, the proposed evaluation concept should attempt to test whether our system can help the user with this process. Hence a task should be defined in which the user is encouraged to find specific sites with the help of the surrounding structures and distinguishable landmarks. As practised in the conventional method, employing landmarks to recognise the current spatial state is a routine solution; hence, by using a study that employs landmark recognition, all four assistance methods can be equally and appropriately compared.

For the sake of an accurate estimate of the effects caused by each augmentation method, the user is asked to repeat this localisation task with each assistance category. However, if the targeted site remains the same during all four methods, there would be a training effect that would result in the user performing significantly better in the succeeding paradigms, solely owing to the adaptation to the site location. On the other hand, if only one site is to be located for each method, the chances of this point introducing error into the measurement (e.g. a false positive, a false negative or even an outlier) would significantly grow. As a countermeasure for these challenges, four different point-clusters were devised. These clusters were placed on the anterior face of the kidney's inner structures, near four prominent landmarks (fig. 4.1e), i.e.

- at the upper pole of the kidney (fig. 4.1a),
- in between the upper pole and the hilus, near the renal vessels' entry point (hilus slit)(fig. 4.1b),
- in between the lower pole and the hilus, near the ureter's exit point (hilus rise) (fig. 4.1c), and
- at the lower pole of the kidney (fig.4.1d).

In these clusters, three points were defined in a localisable way (based on the surrounding structures). Each method is tested using one of these point-clusters without repetition. In addition to reducing the errors, the designation of the three target points retains the time strains and limits of the evaluation process.

To avoid any possible bias among and in between the points and methods, all methods and points should be balanced across subjects to ensure equal opportunities provided for each one. In turn, this prevents any unwanted effects, risen from the testing order of display methods, as well as any possible favourable points for a specific display method (regarding difficulty), from affecting the result.

4.1.1 Accuracy

To be able to measure the accuracy of our methods, first, the definition of this term should be clarified. In this case, the accuracy is derived from the Euclidean distance between two points in

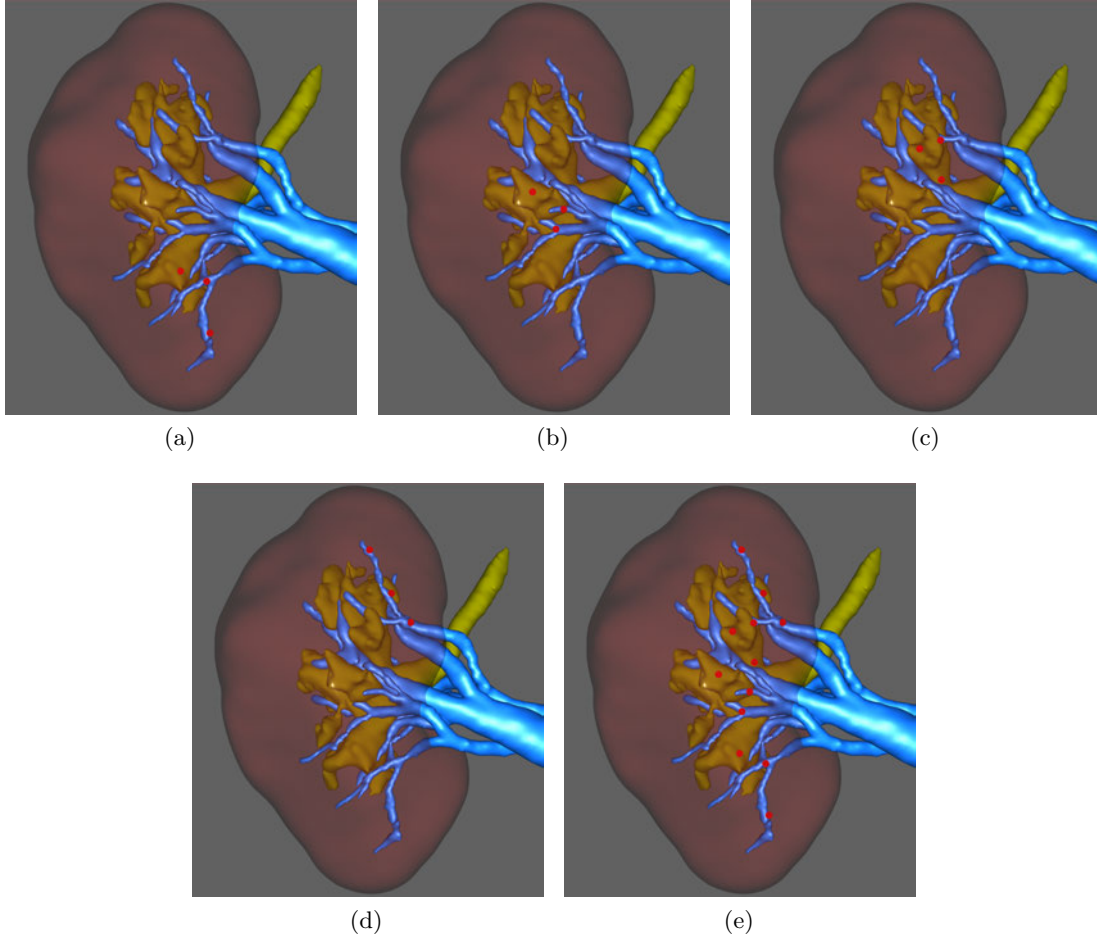


Figure 4.1: Target Point-clusters on the Kidney.

The points placed a) at the upper pole, b) in between the upper pole and the hilus, c) in between the lower pole and the hilus, d) at the lower pole, and e) in total.

3D space. The first point is the target which the user is tasked with locating. While the second point is the position at or the direction in which the user is pointing. Intuitively, with a lesser distance between the two points, the reported accuracy would be higher.

As depicted in figure 4.2, two distance measurements were proposed as possible assessment tools for this definition of accuracy,

- calculation of the distance d_p between the target point and the pointed position on the mesh, and
- calculation of the distance d_d between the target and the pointing direction (closest distance between a point and a line).

In the first technique, after pointing at the kidney mesh, the collision point between the pointing sphere and the mesh is recorded. As mentioned earlier in section 3.3.2, due to the sphere format of this collision criterion, the errors and possible jitters in the result is significantly reduced. If the pointed position lies beyond the penetration depth of our pointing sphere, the

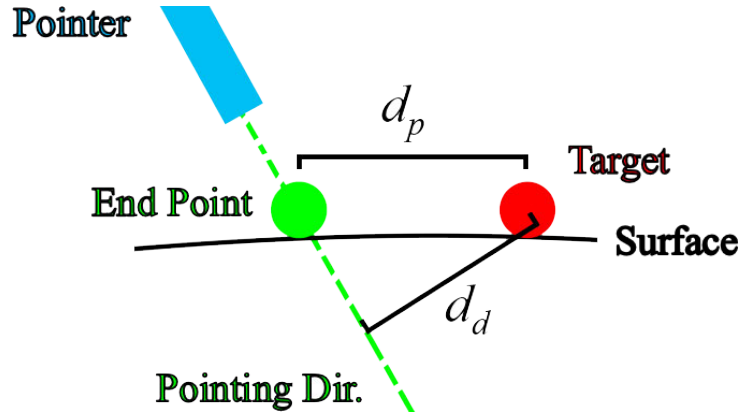


Figure 4.2: Definition of Distance Measurements in the Evaluation.

distance between the needle tip and the target will be calculated instead. The main reason behind this design was to prevent the system from reporting huge outliers in case of small jitters and error in the registration process. However, with both of these failsafe measures in place, the probability of reporting false or exaggerated values (positive or negative) are still of considerable amount. As an example of this outcome, one can give the instance in which the pointed position is a place inside the kidney where the mesh (due to registration errors, jitters, and spherical casting) is on the borderline of our penetration depth. In this case, the possibility of reporting a substantial false negative value would be noticeably increased.

Alternatively, in the second technique, an infinite line is drawn in the direction of pointing. Here, instead of calculating the distance between two points, the closest distance between the target point and this line is calculated. Even though this method is not without errors, it is less prone to reporting false and exaggerated values. In the previous scenario, due to the endpoint not being bound to the mesh, this method can report a much smaller variance between the actual pointed position and the resulted pointed position. It should be noted that this technique can also report a false value. For instance, supposing the target point is beneath the pointing device, while the user is pointing to another position, further away from the target point. In this case, the measured distance could be reported as a false positive. Due to the distance between the pointing device and the inner structures of the kidney, this error value will not be very considerable.

4.1.2 Task Completion Time

Another effect that this study aims to investigate is that of our display methods on the duration of the RSRP. To investigate this duration, one could measure the time spent from the beginning to the end of the task. Every respective interaction performed by the user to achieve the goal of the task should be considered in this duration. Accordingly, in each trial, the duration between the appearance of the target point and completion announcement by the user defines the time value. During this time, the user can make any interactions with the system, which they find necessary (i.e. exploring the kidney visualisation).

4.1.3 Workload

While performing a task, every human has certain limits, e.g. mental, physical, skills, and time, that needs to be regulated in order for the task to be accomplished; however, this is not always the case. Any of these limits may be under or over-burdened at points during the task. The extent to which these limits are pushed during a task can be defined as the respective workload. As it is deductible from the definition of workload, measurement and analysis of a task can prove vital in broadening the understanding of its aspects. Additionally, one can also investigate whether the operation was performed in the most efficient way possible, or some changes are required. While over-burdened workload could be exhausting and overwhelming for the user and lead to failure, the under-burdened workload could pose the same threat by causing lack of focus and motivation for the user to complete the task. [148,149]

In the following, the most established workload measurement techniques are mentioned [148].

- Performance measures, which would take into account only the primary task being performed or the output of the operator. The instances of these measures are:
 - measuring speed and accuracy,
 - measuring activity, and
 - task analysis.
- Indirect measures, in which a secondary task will be introduced, and the performance of the user on this task is measured. These tasks can be of visual, mental, psychomotor, or mixed nature.
- Subjective measures, that measure the workload by asking the users about their experience regarding the performed task. These measures can be separated into two groups:
 - qualitative techniques, and
 - quantitative techniques.
- Physiological measures, where the workload is measured by recording physiological phenomena, e.g. heart rate, Pupillary response, and metabolism, via external sensors.

Given our current case, using the performance measures to obtain workload is not particularly useful. The speed and accuracy of the participant are already being recorded and analysed, each on a different criterion. As for activity, the variability of performance and interactions of the user with different display methods cannot be observed in the proper granularity that could provide useful information. Finally, Task analysis could prove insufficient in distinguishing the source of workload, since the steps of the task in all methods have the same enumeration.

As for physiological measures, even though they are great sources of measurement with high reliability and correctness, using them requires external tools and integrations which are often expensive and hard to acquire. Since these instruments were not available for this thesis, and

the implementations could exceed the proportions of this thesis, physiological measures were also disregarded.

Indirect measures provide more information on the current state of the operator's workload capacity, which can prove beneficial, given our situation. In designing a secondary task, one must always consider the directionality of the secondary task's workload regarding the primary task's workload. If these tasks are wholly coinciding or unrelated, this can result in a wrong measurement. Another disadvantage of the indirect measures is that the user might often disregard the primary task to accomplish the secondary task, which could cause an unreliable measurement in both the workload and the task performance.

As discussed by *Casner et al.* and *Rubio et al.* [148, 150], Using subjective measures could add more dimensions to our assessment of workload. In subjective methods, the user usually gives ratings to different dimensions of workload, e.g. mental, physical, and temporal. Despite the eligibility of quantitative measurement techniques, having a qualitative value could ease the process of analysis and comparison of our display methods. On this account, choosing a subjective quantitative measurement appears as an appropriate choice. There are various subjective quantitative workload tests available, the most common of which are,

- Subjective Workload Assessment Technique (SWAT),
- National Aeronautics and Space Administration Task Load Index (NASA-TLX), and
- Workload Profile (WP).

It should be noted that these tests are not without disadvantages. Most dominant of these disadvantages are by-products of the nature of the assessment itself, that is subjectiveness. This characteristic implies that each person, who takes this test, has a unique perspective on the definition of the scale grains, dimensions, and workload. Another noticeable problem with these tests is the absence of real-time measurement. Since in most cases, the user attends the test after completion of the task (or series of tasks), (s)he would only recall the first and final impression of the system. Finally, given a situation where multiple methods are to be evaluated, if all the methods were to be evaluated with only one test, the result would mainly be based on an impression of the final method that the user has experienced.

Although these measurement errors can not be entirely omitted, by explicitly instructing the user and clarifying the goals and definition of the test, they can be considerably decreased. For instance, the previously mentioned issues could possibly be solved through illustrating each dimension of the test, measure the workload after each method, and explicitly advise to only rate the most recent experience without any comparison.

Due to the multisensory nature of our system, implementation of a secondary task as well as providing an ideal condition in which it can be effectively perceived is exceptionally challenging and error-prone. Hence, the subjective measurement method was selected. Considering all the possible subjective measurement options, on account of its multidimensionality, intuitiveness

and individualism for each user and method, NASA-TLX was chosen. Similar studies, such as [99, 115, 122], also implemented NASA-TLX as their workload measurement technique.

NASA-TLX is practised through a questionnaire [151]. In this questionnaire, six dimensions are defined, i.e. mental demand, physical demand, temporal demand, performance, effort, and frustration. Each one of these dimensions is scored using a scale with 21 tick marks. The user marks the tick that (s)he deems as a suitable representation of the experienced workload. To calculate the raw TLX score TLX_{RS} , the following formula can be used,

$$TLX_{RS} = \sum_{i=1}^6 (Score_{Ri}), \quad (4.1)$$

Where i is the respective workload dimension, and $Score_{Ri}$ is the raw TLX score marked by the user. The raw TLX rating TLX_{RR} can also be computed using the following formula,

$$TLX_{RR} = \frac{\sum_{i=1}^6 (Score_{Ri})}{6}, \quad (4.2)$$

Where the raw TLX score will be normalised on a scale between 1 and 21. Alternatively, 0 and 20 is the other scaling range, however in this study 1 and 21 is selected.

While using raw TLX, all the dimensions are weighted equally. In other words, it is assumed that the workload on each dimension carries an identical consequence, regardless of the task.

The alternative to this method is the weighted TLX score, where participants are asked to assign a weight for each dimension based on the relevance that they felt to their performed task [151]. This comparison is accomplished via 15 paired dimensions, from which only one should be selected. Eventually, this leads to a weight between zero and five in each dimension, and a total weight of 15 distributed among six dimensions. Weighted TLX score TLX_{WS} can be computed using this formula,

$$TLX_{WS} = \sum_{i=1}^6 (Score_{Ri} \times Weight_i), \quad (4.3)$$

Where $Weight_i$ is the weight assigned by the user to workload dimension i . Weighted TLX rating TLX_{WR} can also be calculated with the following formula which normalises the score between 1 and 21,

$$TLX_{WR} = \frac{\sum_{i=1}^6 (Score_{Ri} \times Weight_i)}{15} \quad (4.4)$$

The main advantage of using weighted TLX score compared to raw TLX score is that it reports a unique score for each user. Ultimately, it can supply a more accurate estimation of the workload

that the user has experienced in each task [148]. Hence, the weighted TLX score/rating is selected as the default workload measurement tool for this evaluation. The NASA-TLX questionnaire used in this study is provided in the appendix with figure A.2 for the dimensional scales, and figure A.3 for the 15 pairs of dimensional comparison.

4.1.4 Design Summary

Following the arguments presented earlier, the final design of the evaluation method is proposed. In this design, a localisation task of four point-clusters, one cluster per an assistance method, is considered. Three points were devised per a point-cluster (a total of 12 points as depicted in fig. 4.1e). The participant is explicitly told to locate these points as accurately and quickly as possible, using the landmarks and structures that are recognisable. The duration of each trial is recorded from the appearance of the target point until the concluding verbal cue of the user. For the accuracy, the distance between the target point and the pointing direction is measured. Finally, for the workload, NASA-TLX questionnaire is used to acquire weighted TLX ratings, once per an assistance method for each participant.

In the ideal scenario, to administer a completely balanced study, one should probe all possible permutations of four assistance categories (i.e. $4 \times 3 \times 2 \times 1 = 24$ unique cases), where every user would experience a unique progression in display methods. Eventually, the number of these instances can be reduced, supposing that every method is repeated equally in each progression slot.

4.2 Setup and Settings

The setup used for this solution, and its evaluation is depicted in figure 4.3. This setup consists of seven main components (fig. 4.4),

- an optical tracking system (4.4a),
- a 3D printed kidney model (4.4b),
- a box with entry ports, replicating a surgical site (4.4c),
- a laparoscopic camera system (4.4d),
- a surgical grasper (4.4e),
- a monitor with an interactive 3D visualisation of the kidney (4.4f),
- a monitor with the laparoscopic video stream (4.4g), and
- a loudspeaker to provide auditory feedback (4.4h).

The optical tracking system (mentioned in section 3.3) is tasked with consistently tracking the position of the 3D printed kidney, laparoscopic camera, and surgical grasper. The 3D printed kidney model is the primary interaction unit between the virtual and real world. This model is

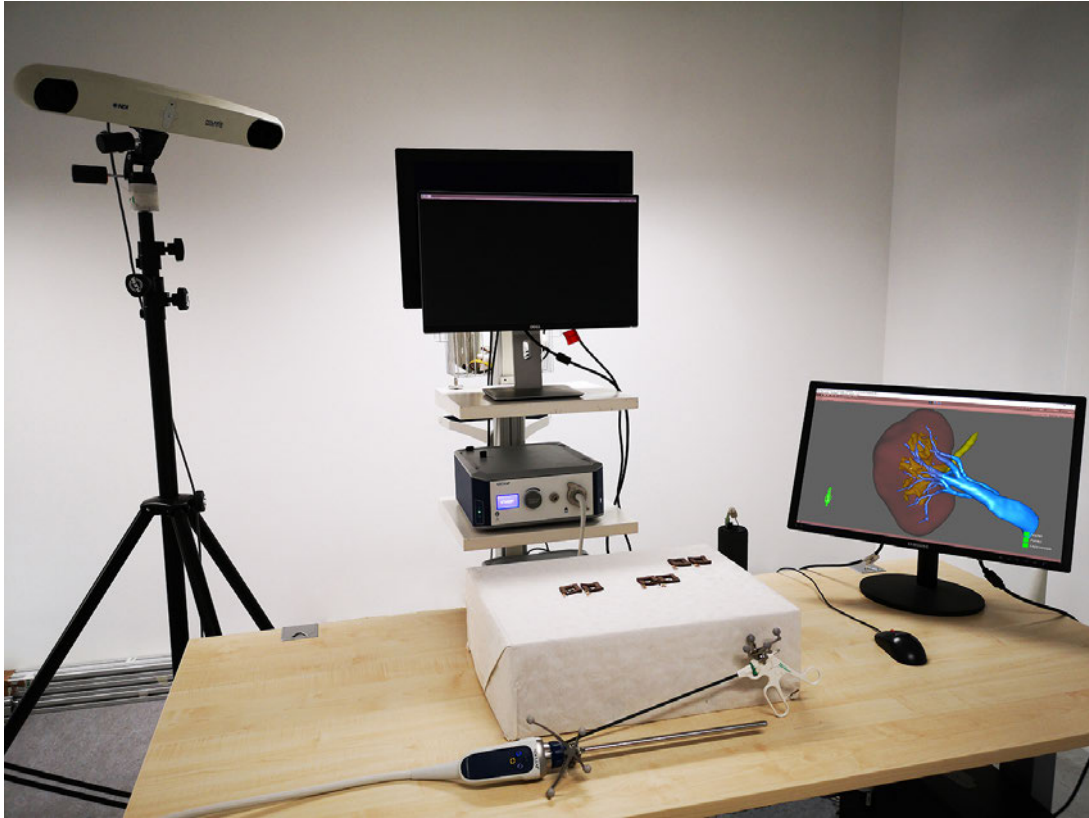


Figure 4.3: Evaluation Setup Overview.

fixated and attached to a set of fiducial markers and is considered as the origin of our virtual coordinates system. The user investigates the renal substructures through this model by exploring its surface and localising the target points.

To simulate a laparoscopic surgical procedure, as well as preventing the user from directly viewing this model, a box with six ports covers the kidney. The participants are explicitly instructed to insert the camera in port one and the surgical grasper in port three of the box, and unless an absolute necessity arises, they should not change the grasper's port.

The laparoscopic camera is used to obtain a live video stream of the events that occur under this box. Starting with the trial, this video stream is viewed in a dedicated monitor. The surgical grasper is used as a haptic and pointing device, with which the user is encouraged to investigate the surface of the model and the targets. The VD overlay is also placed at the tip of the grasper.

The 3D interactive kidney visualisation is assisting the user by replicating the kidney and its structures. This visualisation is similar to the mental image and spatial information that an experienced surgeon has. This model can be moved, rotated, and zoomed, on-demand, in order to provide the user with reliable spatial and structural information. Finally, a loudspeaker will provide the auditory feedback produced by the AD.

The adjustability of the setup was mentioned earlier in section 3.3, yet to achieve a universal measurement and comparable results for our setup, these settings should be set equally for all

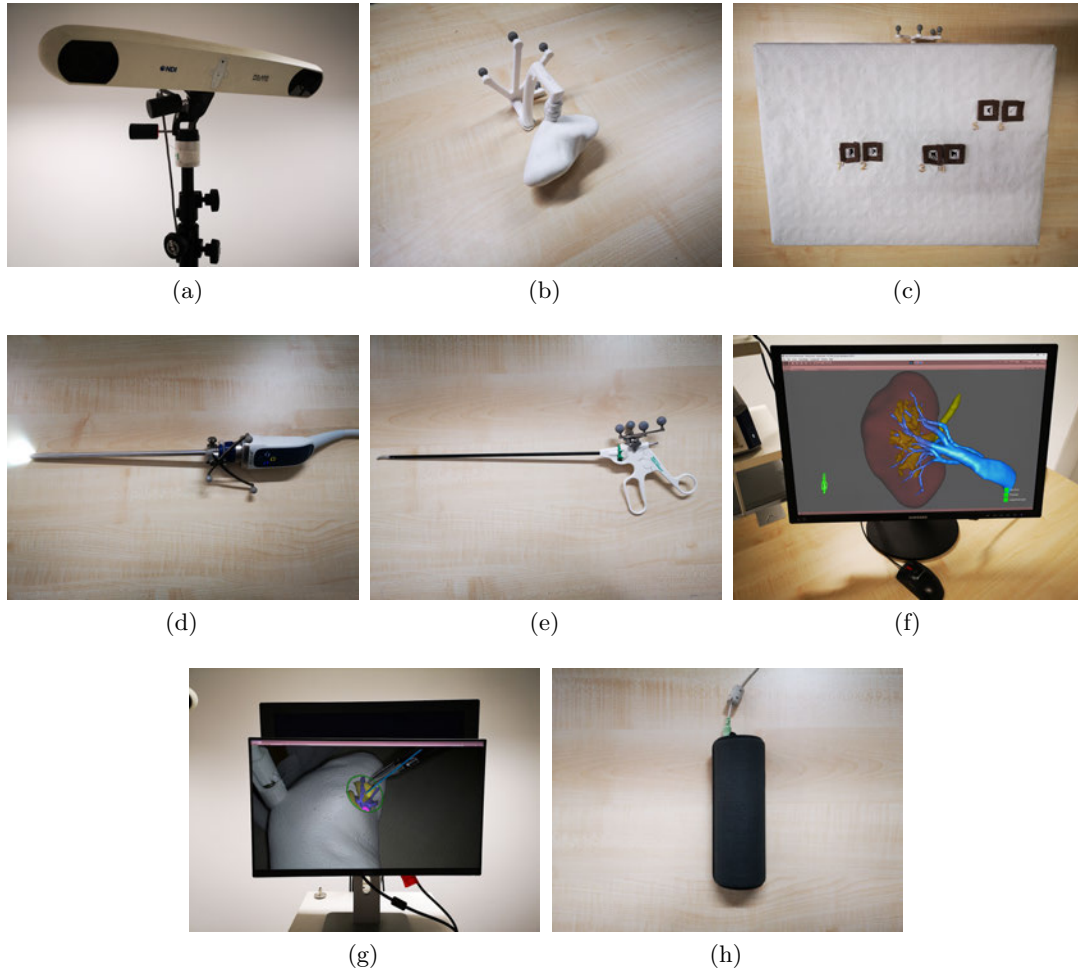


Figure 4.4: Evaluation Setup Components.

These components are a) NDI tracking system, b) 3D printed kidney model, c) the covering box, d) laparoscopic camera, e) surgical grasper, f) kidney visualisation monitor, g) laparoscopic video monitor, and h) loud speaker.

participants. Therefore, based on trial-and-error as well as the current state of our setup, these settings were selected as default,

- a mesh with a size of 160×160 vertices,
- a visual display overlay with a diameter of 20mm ($64 \text{ vertices}/\text{mm}^2$),
- a probing depth of 7mm to 40mm ,
- an auditory feedback response site with a diameter of 4mm ,
- a minimum of 3000 vessel vertices to be reported as a large vessel,
- a minimum of 4500 urinary vertices to be reported as a large urinary structure,
- a vessel colour gradient scale (0 – 100%) with RGBA codes (0, 0, 52, 0) at 0%, (0, 0, 102, 204) at 75%, and (255, 0, 255, 255) at 100%, and

- a urinary colour gradient scale (0 – 100%) with RGBA codes (52, 52, 0, 0) at 0%, (102, 102, 0, 204) at 75%, and (255, 255, 0, 255) at 100%.

The colour selection process is described in details in section 3.2.2.3.

4.3 Population

The suggested solution of this thesis is intended for clinical situations, where surgeons could benefit from these assistance methods, and possibly reduce errors and complications. Hence the ideal population for this evaluation should consist of medical doctors and surgeons who have experience in the field of laparoscopic surgeries. Unfortunately, acquiring this population is arduous and often unfeasible. Alternatively, the closest available population to this demography, are medical students who have some fundamental knowledge and training regarding medical procedures and human anatomy. Thus, medical students were chosen as an immediate target group of this study.

Preliminary to the main evaluation study, a pilot study with four non-medical participants, one male and three females, who reported no previous experience with neither laparoscopic surgeries nor AR, was conducted. The main goal of this study was to find possible shortcomings and implement improvements for the main study, if necessary.

Eleven participants have taken part in the main study, comprising six females, five males, and all right-handed. The participants were in the range of sixth to tenth semester, with the mode and median of 10 and mean of 9.18. The population's age had an average of 26, ranging from 24 to 33, with 25 as the mode and median.

Eight participants reported that they consider themselves musical, where only five of them are currently engaged in playing a musical instrument. Four of the participants had previous experience with laparoscopic setup, with a duration of 1 to 5 hours, the average and median of 3 hours. As for prior experience with augmented reality, seven users reported a range of 1 to 15 hours, with an average of 4.93, two modes of 1 and 2, and a median of 2 hours.

Only one participant reported an uncorrected vision/eyesight, which was a case of myopia that would not affect the engagement and interactions of the mentioned user with the current study setup.

4.4 Study Procedure

Before engaging with the setup, the user fills a participation consent form, a demographic questionnaire, and receives written instructions on what to do and expect from the setup. Afterwards, the user starts the study by going through a tutorial for approximately 15 minutes. During this tutorial, the user is thoroughly instructed on how to handle the instruments and the kidney visualisation. Furthermore, the visualisation and sonification of the structures are demonstrated for the user. In the next part of the tutorial, the user completes a practice round

on the kidney model to get familiar with the actual task of the study. In the final step of the tutorial, the remainder of the time is dedicated to exploring the setup, display methods, and the kidney model. During the tutorial, a red dot provides feedback on the location at which the user is pointing.

During the main trials and for each assistance method, the user goes through three localisation tasks. Before trial initiation, instruction regarding the following assistance method is provided. When the instruction has been read, the user moves to the setup table. Then, (s)he sets the kidney visualisation to a preferred viewing angle and inserts the instruments into the box with the camera in port one and the surgical grasper in port three. At this point, the user has no laparoscopic stream nor the target point on the kidney visualisation. After the user has announced readiness, the localisation task begins with a countdown by the experimenter. As soon as the countdown ends, the laparoscopic stream and the target point on the kidney visualisation appear simultaneously, and the stopwatch begins recording. When the target is located with certainty, the user announces it, and the experimenter stops the task; the resulted time and the distance of the task is logged. At this point, the target point and laparoscopic stream disappear. The same steps are repeated for two other points in the cluster for the same assistance method.

When all three localisation tasks for the same condition are completed, the user is asked to fill the NASA-TLX form. (S)He is explicitly asked only to report the workload that has been experienced during the last assistance method, without any comparison. Afterwards, these steps are repeated for the next methods until all four assistance conditions are tested.

As discussed earlier, the display methods and points were balanced for a more accurate measurement. Accordingly, 12 study scenarios were generated for 12 participants; however, since only 11 participants took part, one of the scenarios was not used in the study. The complete list of these scenarios is provided in the appendix (table A.2).

4.5 Analysis and Results

4.5.1 Descriptive Statistics

As the next step in the evaluation of the proposed method, the results of the user study were analysed. To have a general outlook on the results of the study, first, the descriptive analysis of the current samples was conducted. The results for each dependant variable across all users were then averaged and the standard deviation (SD) was calculated. Afterwards, according to the empirical rule, any trial with a value below ($mean - 3SD$) and above ($mean + 3SD$), for any of the dependant variables, were omitted from the next steps of the analysis process. The results after this, including mean, SD, and standard error of the mean, (SEM) and their box plots (fig. 4.5) is presented in the following. On the other hand, the table and scatter plots of this analysis can be found in A.1 and A.4, available in the appendix.

The distance between the target and direction of pointing for NoA setting lies between $6.027mm$ and $14.998mm$, and has a mean of $11.131mm$ ($SD = 2.817$, $SEM = 0.849$). As for our second

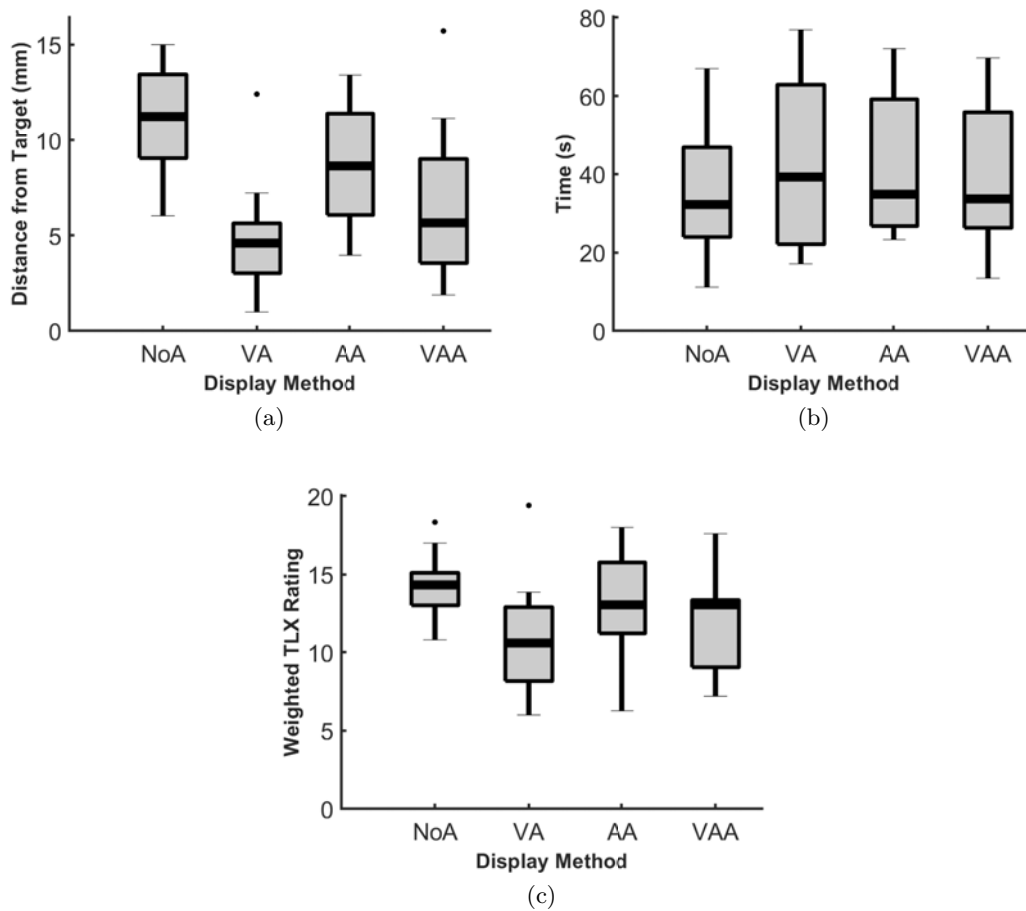


Figure 4.5: Descriptive Statistics through Box Plots.

Box plots of a) Distance from the target, b) Time, and c) weighted TLX rating is shown for each assistance method.

variable, the task completion time, the outcome of the analysis shows a mean of $35.391s$ (SD = 17.019 , SEM = 5.131) with the values placed between $11.169s$ and $66.980s$. Finally, the weighted TLX rating for this method is reported in the range of 10.8 to 18.333 , with an average of 14.139 (SD = 2.182 , SEM = 0.658).

For VA method, the distance shows a mean of $4.801mm$ (SD = 3.153 , SEM = 0.951), and the values ranging from $0.992mm$ to $12.415mm$. The duration varies between $17.117s$ to $76.848s$ with a mean of $42.05s$ (SD = 22.566 , SEM = 6.804). The reported workload rating in this method has a minimum of 6.00 and a maximum of 19.40 , with an average of 10.873 (SD = 3.861 , SEM = 1.164).

On the other hand, the outcome of AA setup, reports a minimum of $3.963mm$ and a maximum of $13.412mm$, and a mean of $8.594mm$ (SD = 3.329 , SEM = 1.004) for the distance. The measured time is a value between $23.290s$ and $72.036s$, and a mean of $41.769s$ (SD = 17.786 , SEM = 5.363). The workload rating of this method is in average 12.927 (SD = 3.462 , SEM = 1.044) with the lowest value of 6.267 and the highest value of 18.00 .

Finally, for VAA solution, the distance values are varying from $1.884mm$ to $15.712mm$, with a mean of $6.432mm$ ($SD = 4.230$, $SEM = 1.275$). The average duration of a trial using this method is $39.833s$ ($SD = 18.285$, $SEM = 5.513$) with the maximum value of $69.696s$ and the minimum value of $13.437s$. As for the workload rating, VAA scores an average of 11.927 ($SD = 3.296$, $SEM = 0.994$) with a minimum rating of 7.20 and a maximum rating of 17.60 .

At first glance, boxplots show an improvement in accuracy using any of the proposed AR assistants in comparison to the absence of AR. These changes are not as visible in the time domain, implying our methods do not have much of an improving impact on the duration of the task. On the other hand, one can note the reduction of workload rating for all the AR methods in contrast to NoA method. Nonetheless, to investigate the validity of these observations, further tests must be conducted.

Under the influence of AR methods, compared to the control condition, the average accuracy of localising task has noticeably increased. While AA was able to increase the accuracy by 22.79% , VA improved the accuracy by 56.86% . The combination of two methods, namely VAA, was able to achieve a 42.21% improvement. The workload rating was also improved in all AR cases, i.e. VA with 23.09% , AA 8.57% , and VAA 15.64% . These improvements suggest that an AR assistant can improve the accuracy and workload aspect of the procedure. On the other hand, there is no constructive effect on the duration of the procedure; VA increased the duration by 18.81% , AA extended the time by 18.02% , and VAA decreased the completion speed by 12.55% .

4.5.2 Analysis of Variance

To investigate the differences and effects of the four assistance methods, analysis of variance (ANOVA) was selected as the preferred tool. Since the study is based on a two-by-two factorial design, it is appropriate to choose a two-way ANOVA method. As the final criterion for the selection of our analysis method, one should point out that each participant has attended all conditions, i.e. all four assistance methods. Given this final argument, a suitable choice for our analysis method is the two-way repeated measures ANOVA (RMA). The results of these analyses for the accuracy, time, and workload are presented in table 4.2.

According to *Field et al.* [152, 153], to estimate the effect size of each independent variable, three methods are available, i.e. eta-squared (η^2), partial eta-squared (η_p^2), and omega-squared (ω^2). η^2 often tends to be biased, specifically in small samples sizes, where it also tends to exaggerate the effect size. η_p^2 is a variant of this test to reduce this exaggeration; however, the bias is still apparent. Eventually, the ω^2 offers a more reasonable solution with no bias, as well as impartiality towards sample size. The effect size based on the ω^2 test results can be interpreted as,

- $0.01 \leq \omega^2 \leq 0.06$, as small,
- $0.06 \leq \omega^2 \leq 0.14$, as medium, and
- $0.14 \leq \omega^2$, as large.

| RMA Results for Accuracy | | | | | | |
|--------------------------|----------------|----|-------------|--------|---------|------------|
| Cases | Sum of Squares | df | Mean Square | F | p | ω^2 |
| VD | 198.292 | 1 | 198.292 | 17.413 | **0.002 | 0.379 |
| Residuals | 113.874 | 10 | 11.387 | | | |
| AD | 2.259 | 1 | 2.259 | 0.285 | 0.605 | 0.000 |
| Residuals | 79.366 | 10 | 7.937 | | | |
| VAD | 121.440 | 1 | 121.440 | 11.239 | **0.007 | 0.280 |
| Residuals | 79.366 | 10 | 7.937 | | | |
| VD * AD | 47.756 | 1 | 47.756 | 4.301 | 0.065 | 0.108 |
| Residuals | 111.026 | 10 | 11.103 | | | |

| RMA Results for Time | | | | | | |
|----------------------|----------------|----|-------------|-------|-------|------------|
| Cases | Sum of Squares | df | Mean Square | F | p | ω^2 |
| VD | 61.360 | 1 | 61.360 | 0.323 | 0.583 | 0.000 |
| Residuals | 1902.316 | 10 | 190.232 | | | |
| AD | 47.621 | 1 | 47.621 | 0.236 | 0.638 | 0.000 |
| Residuals | 2021.494 | 10 | 202.149 | | | |
| VAD | 108.547 | 1 | 108.547 | 0.496 | 0.497 | 0.000 |
| Residuals | 79.366 | 10 | 7.937 | | | |
| VD * AD | 203.212 | 1 | 203.212 | 2.056 | 0.182 | 0.009 |
| Residuals | 988.492 | 10 | 98.849 | | | |

| RMA Results for Workload | | | | | | |
|--------------------------|----------------|----|-------------|-------|--------|------------|
| Cases | Sum of Squares | df | Mean Square | F | p | ω^2 |
| VD | 50.062 | 1 | 50.062 | 6.345 | *0.030 | 0.122 |
| Residuals | 78.898 | 10 | 7.890 | | | |
| AD | 0.068 | 1 | 0.068 | 0.013 | 0.911 | 0.000 |
| Residuals | 52.078 | 10 | 5.208 | | | |
| VAD | 26.194 | 1 | 26.914 | 4.149 | 0.069 | 0.106 |
| Residuals | 79.366 | 10 | 7.937 | | | |
| VD * AD | 14.128 | 1 | 14.128 | 1.468 | 0.253 | 0.014 |
| Residuals | 96.217 | 10 | 9.622 | | | |

Note. Type III Sum of Squares

* $p < .05$, ** $p < .01$, *** $p < .001$

Small $0.01 \leq \omega^2$, Medium $0.06 \leq \omega^2$, Large $0.14 \leq \omega^2$

Table 4.2: RMA Results for within Subjects Effects.

On account of the study factors being two-level independent variables, the sphericity tests and post-hoc test are not necessary [152, 153]. The effects of visio-auditory display (VAD) is a direct comparison of the control condition (NoA) and VAA.

Based on the values from table 4.2, for the accuracy, a statistically significant effect of VD, $F(1, 10) = 17.413$, $p = 0.002$, and a large effect size of $\omega^2 = 0.379$ was observed, thus rejecting the $H_{V_{a0}}$. On the contrary, there is no statistically significant effect of AD ($p = 0.605$) nor an

interaction between AD and VD ($p = 0.065$), regarding the accuracy of the method; failing to reject H_{Aa0} and H_{AVa0} . However, in our population, an improvement in accuracy can be detected whenever AD is supplemented with VD, which can hint towards a possible trend. Nevertheless, this outcome is uncertain unless significant evidence is provided with follow-up experiments. Even though not part of our NHST hypotheses, VAD shows a significant improvement in accuracy, $F(1, 10) = 17.413$, $p = 0.007$.

Concerning the time, neither VD ($p = 0.583$), AD ($p = 0.638$), nor their interaction ($p = 0.182$) shows a statistically significant effect, unable to reject H_{Vt0} , H_{At0} , and H_{AVt0} . Ultimately, one can loosely deduce that all three AR methods can be accomplished in the same or, with more training, less amount of time than NoA, while offering higher accuracy. Once more, it should be noted that these interpretations require more experimentation and statistical evidence in order to be considered valid.

As for the final variable, VD possesses a statistically significant effect on the workload rating, $F(1, 10) = 6.345$, $p = 0.030$, with a medium effect size of $\omega^2 = 0.122$, hence rejecting the null hypothesis H_{Vw0} . While AD ($p = 0.911$), and the interaction of AD and VD ($p = 0.253$), demonstrating no significant effect; hence not rejecting H_{Aw0} and H_{AVw0} .

To find the directionality of both the accuracy and workload rating, improved by VD, one can compare the mean of each level using figure 4.6, without running a post-hoc test. With the introduction of VD, the accuracy is increased (the distance is decreased). As for workload rating, it is also visible that VD reduces the workload experienced by the user. These effects on accuracy and workload are mainly expected; since instead of performing all the mental and

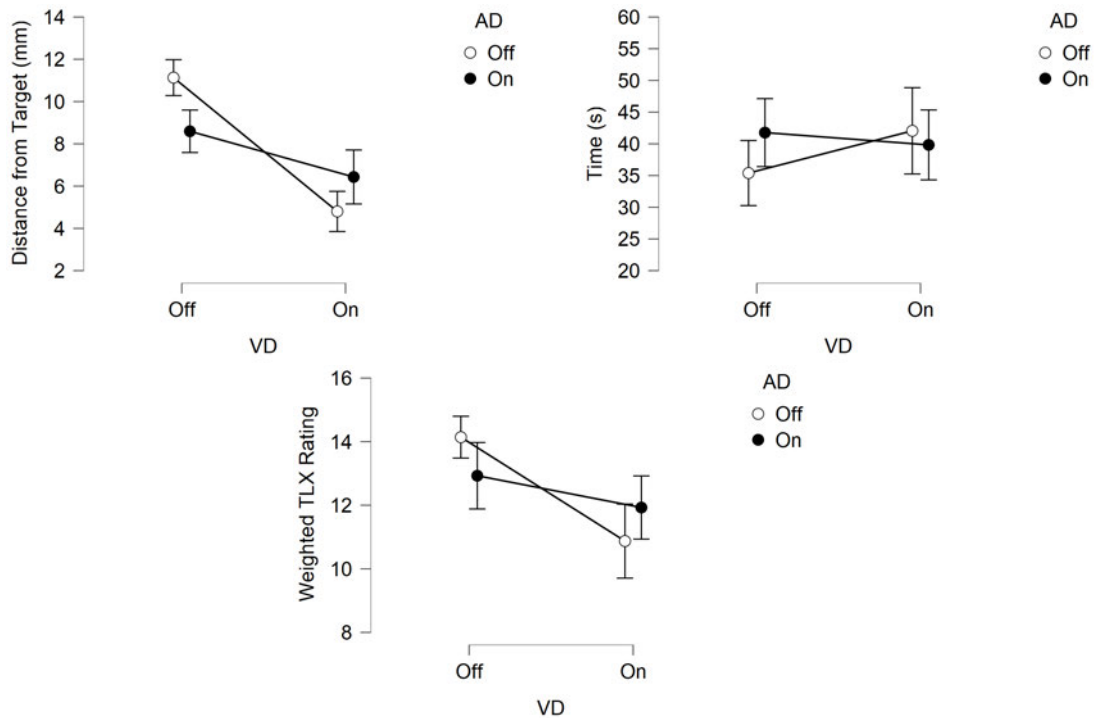


Figure 4.6: Descriptive Plots for RMA Results.

physical demands of locating the point, without having much information on inner structures of the kidney, the users are provided with this information.

4.6 Evaluation of the Requirements

The final step in our evaluation process brings us back to the requirements mentioned in section 3.1; where a series of functional and non-functional requirements were set for the solution. It is now essential to confirm whether the implemented solution fulfilled these prerequisites. An overview of these requirements and their state of fulfilment can be found in table 4.3.

The functional requirements are the pillars of the solution; hence fulfilling them is a necessity. All the designs in this project were devised with respect to these requirements. As for their counterparts, i.e. non-functional requirements, they were mostly achieved. Due to the constraints and proportions of the project, satisfaction and registration errors were not thoroughly investigated. Nevertheless, the final solution has fulfilled these standards predominantly.

| Requirement | Status | Note | |
|-----------------------------|------------------------------|---------------------|--|
| Functional Requirements | Component Tracking | Fulfilled | Through optical tracking system. |
| | Augmented Reality | Fulfilled | Using visualisation and sonification to relay the information. |
| | Registration | Mostly Fulfilled | The AR contents were successfully registered; however the error is unknown. |
| | Recognition and Localisation | Mostly Fulfilled | All AR methods show improvements in localisation accuracy in the evaluation population (VD significant). |
| Non-functional Requirements | Satisfaction | Partially Fulfilled | In a post-evaluation informal interview, more than half of the users were happy with the AR methods, though registration errors, AD design, and spatial mapping were of concern to others. |
| | Efficiency | Mostly Fulfilled | All AR methods show improvement in workload (VD significantly), but a small increase in the duration of the task. |
| | Effectiveness | Fulfilled | All AR methods show improvement in accuracy (VD significantly). |
| | Display Frame Rate | Fulfilled | An average of 28 FPS (24 - 32) were observed while interacting with the setup. |

Continued

Continued

| | | |
|---------------|-----------|---|
| Latency | Fulfilled | The information refresh rate is 60 FPS, while our system holds, on average, 28 FPS, ensuring the availability of the information at all frames. |
| Robustness | Fulfilled | No systematic failure or errors occurred while engaging with the setup. |
| Minimalistic | Fulfilled | The design is transparent and occupies the least possible amount of space on the viewport. Furthermore, the size of this feature is adjustable. |
| Colour-coding | Fulfilled | This topic is discussed thoroughly in 3.2.2.3. |
| Auditory Cues | Fulfilled | Discussed in 3.2.3. |

Table 4.3: Evaluation of Requirements.

5 Discussion

5.1 Outcome

After evaluating the methods proposed by this thesis, it is now possible to have a thorough discussion about the results. Even though only two of our nine null hypotheses were rejected, the results are not altogether unexpected. In the following, these outcomes are discussed in more detail. While VD shows significant effect on accuracy and workload, and VAD on accuracy, the following discussion is majorly based on all the visible results in the current study.

While AR assistants are present, the average localisation accuracy was improved. This improving effect is also visible in the workload rating across all AR assistants, compared to the control condition. However, no positive impact on the task completion duration was noticed among these evaluation conditions. Whether visual, auditory or both, these results have already been observed in other studies as well [12, 98, 118, 122].

It should be pointed out that laparoscopic tasks are heavily dominated by vision. Hence, being able to see the investigation site and its sub-surface structures directly can significantly impact the localisation result. In this case, having two separate colour gradients for each type of sub-structure, being able to observe internal and external landmarks, and distance-coded shading could be recognised as the main contributors. This improvement in the interpretation of the information reflects itself in the workload rating of the task (in VA), in which the users reported the lowest experienced workload, on average. This reduction in workload can support the idea that due to the visual nature of the laparoscopic task, a unisensory input could grant the user a higher level of focus and attention, in addition to fluent information interpretation.

Furthermore, one can reason that the prolonged duration of the task during VD is the by-product of focusing on increasing accuracy. As reported orally by some users and observed during the trials, when provided with the visualisation of the kidney's content, they are determined to reduce the localisation error as much as possible. This fixation with high accuracy prompts the user into neglecting the time constraint. However, in real surgical scenarios, the surgeons might react differently; meaning they might disregard the accuracy in order to reduce the length of the operation. For instance, if the operation is prolonged already, and ischaemia time is increasing, the surgeon might decide to compromise the accuracy for the sake of the ischaemia time.

The VAD method also shows a significant boost to accuracy. As before, this substantial improvement could be associated with visual feedback rather than auditory. It can further be argued that the decrease in the accuracy is due to the introduction of the auditory method. While some studies show that cross-modal effects can improve attention spaces, there is an equal

number of cases where a reduction in attention as a result of multisensory input stretching the attention span too thin, is demonstrated [154–158]; nonetheless, this effect is heavily influenced by the prescribed task. The reminiscence of this effect is further visible in VAA, where both display methods accompany each other. We already have established that the VD significantly improves the accuracy; however, this improvement decreases as we introduce the AD. This result can suggest that the presence of the auditory system in such a visually dominant task, can be recognised as a source of frustration and distraction; hence the longer task completion duration and the lower accuracy scores. The workload rating also increases in comparison to VD alone; further supporting the idea of the auditory feedback putting strains on the user.

As stated in chapter 1 of this study, the presented auditory method is not aiming to replace the visual display, but to reduce the paramount reliance and sensory overload of the visual domain. It is only natural that in a (mainly) visual task, an auditory method would not be as effective as a visual solution.

Although the AD has a smaller effect on the accuracy compared to VD, this increase in the accuracy heralds that the users are, indeed, utilising the display in order to locate the target. However, this difference between the accuracy gains can be explained through the human’s sensory perception. Ample literature has indicated the accuracy and reliability of visual perception in recreating spatial information [159–164]. This effect has also been investigated for the auditory domain but not as exhaustively. Nevertheless, there are established literature explaining how the auditory domain could lack in compensating what visual domain offers [159–170].

While in the visual domain, multiple signal sources can be detected through the direct decoding of neural activity; this process is more complicated when it comes to auditory sources. The process of auditory localisation and interpretation is a series of signal processing based on the physics of the acoustic waves and ear shape; hence, a delay in acquisition and interpretation of auditory signals, in comparison to visual cues (even though largely minute), is an expected outcome [162]. This notion hints that when presented with the same information in both visual and auditory domain, due to the requirements of the hearing process, auditory cues might be neglected, whether consciously or unconsciously. Furthermore, in an average human, the dependency on the visual domain is significantly superior to any other sensory input. According to the results from [159, 164, 166, 167, 169], though humans can localise an event solely through hearing, by introducing a false visual cue, one can explicitly replace the reported location of the source [160]. This phenomenon could further prove the dominant presence of visual cues in human perception. The literature suggests that in multisensory scenarios, visual input is mainly used as a tool to calibrate other senses to improve the accuracy and reduce the spatial uncertainty and errors [162–164, 169, 171].

However, there is more to sensory perception than just the dominance of the visual domain. In cases where visual sensory is not present, e.g. absolute darkness or congenital or acquired blindness, an enormous boost in auditory spatial estimation is observable [159, 164, 166, 167, 169]. Alternatively, with more repetition and practice in the auditory domain, one can achieve a higher degree of understanding and fluent interpretation of the provided auditory methods.

Another potential issue arises from the limitations of auditory channels in receiving and deciphering the information. Auditory channels can successfully distinguish two streams of input simultaneously while only being able to analyse one [60, 128]. On this note, it should be pointed out that the current method is occupying the full width of this channel with two streams being transmitted at all time. A possible strategy in this scenario is to analyse each auditory cue subsequently. Through this strategy, the user selects one of the streams that (s)he is hearing and then investigates the information that is mapped to the acoustic properties. As discussed earlier, due to the intricate procedure of auditory perception, and lack of training, this can prolong the duration of the task even further.

One of the studies that attempted to test the auditory spatial attention was conducted by *Spence et al.* [154]. This study was a pioneer in addressing a key factor in the sensory domain, which is the individually-unique interpretation method of the visual and auditory system. The lack of a proper task definition for each sensory domain was pointed out in this paper, explaining the null and primed effects, as well as mixed outcomes of the previous publications. As a countermeasure, *Spence et al.* designed the same, but specific, paradigms for each of the visual and auditory systems. The results suggested that auditory and visual system are not considerably different in localisation tasks if given the right circumstances.

Currently, our evaluation method does not contain such an equalised approach for testing the auditory system, and this is where the following problem becomes more so apparent. Suppose in a scenario, where the visual stimuli must be recognised through a spatial mapping that is established by the auditory sensory system. In our evaluation method, the user sees a point on the 3D visualisation. To address this spatial information, a mental spatial mapping based on the auditory domain must be formed. Finally, the point is located in real-world using this mental spatial image. Now, this might not be as difficult and complicated if the user was more familiar with the setup and the tones, though this was not the case in our evaluation, due to the limited time constraints. This effect might be completely different if an expert were to participate in the study. The experience and spatial understanding of a laparoscopic surgeon, in contrast to a novice medical student, could immensely contribute to the reliability of the results.

Hansen et al., *Plazak et al.*, and *Bork et al.* reported a significant improvement in the accuracy when using AD [98, 99, 118]. However, it can be argued that this outcome is due to the designs being one dimensional. *Black et al.* stated that the absence of a significant effect on the accuracy could be due to the extended dimensions of auditory feedback [122]. Furthermore, the task described in the studies conducted by *Plazak et al.* and *Bork et al.* is the localisation of a target, where the user receives feedback about this target [98, 99]. In our paradigm, even though a localisation task is performed, the feedback is not target-specific. This highly variable feedback could also be another factor which affects the difficulty of the task.

As hinted in previous sections (4.6 and 3), the sound designs might be another contributor to the shortcomings of AD. A few of the users reported frustration, confusion, and annoyance while using the AD. Though the design aimed for maximum possible contrast between the two sounds,

this aspect might have caused more harm than good. The complete inorganic nature of this duality could affect the attention space of the user; distracting them from focusing on one sound and, eventually, reducing certainty and precision, as well as increasing the time required for deciphering the information. Even though this issue was tried to be addressed through low-pass filters (volume control), it persists. Furthermore, even with all the noise and jitter reduction methods in place, the jitters of tracking system causes uncertain feedback in some scenarios, due to the rapid movement of the tool.

Furthermore, the running water sound is a complex natural sound. As described in the literature, auditory icons are usually brief and to the point. Hence, another possible explanation for the mentioned issue is the water sound itself and its continuous presence in our feedback design. Possibly, noticing and processing the rapid changes in the pressure and wave metaphor can stretch the attention space of the user too thin. On the other hand, Due to everyday connection of us to the water sound, these alterations can shift our attention towards it even to an unconscious level. Additionally, continuously hearing a complex and progressive auditory icon can put strains on one's attention and perception, and even further when there are two simultaneous auditory streams. The constant presence of both sound sources while engaging with the system can burden the user, reducing their focus and attention.

As mentioned earlier, unfortunately, due to the nature of the task and evaluation method, not much can be said with certainty about the AD sound designs and effectiveness. A study must be carried out in which the user can train competently with the AD before drawing any more convincing conclusions.

The nature of the laparoscopic task is highly visual; hence a sensible strategy executed by the users is to rely more on the feedback received from the visual domain. Additionally, this trait imposes design constraints in which an equalised evaluation method is unfeasible. It should be mentioned that we are investigating whether our methods could aid the specific laparoscopic task, and not that if our methods can perform equally well under inorganic scenarios.

5.2 Outlook and Future Work

The work of this thesis is a preliminary prototype of a novel solution to a specific laparoscopic task, i.e. the RSRP during LPN/RAPN; hence it is essential to recognise and highlight the shortcomings of the current methods in order to improve upon them in the future steps of this solution.

This work was chiefly based on two fundamental assumptions; first, the absence of deformation in the kidney during the surgery and after resection, and second being that the resection site is a flat surface and not cavitory. On account of promising results under these assumptions, as a future step, it is highly encouraged to feature curvature (concave or convex) tools in order to investigate non-flat resection sites. In addition, further down the road, a solution with deformation compensation and motion correction can be developed to solve the current commonplace shortcomings of this system.

The VD manages to provide ample information regarding the spatial information and the type of the visible structures, and significantly improve the workload and accuracy. However, it fails to indicate the exact distance from the instrument tip to the structure's surface. In future works, this value can be reported in an optimised and non-distracting way.

In contrast, AD was not able to achieve statistical significance, yet it was able to improve the average accuracy and workload in our population. As discussed earlier, this can be a by-product of human's perception and attention phenomena. While, in theory, the sound design aims to boost the attention and learning rates of the user, these effects were not completely visible in the final results. It is possible that by implementing a less complex tone than water sound, and more discrete and concise feedback, further improvements would be observed. However, a more competent study must be conducted to ensure sound design flaws of this method.

As for the structure of the display, few alternatives can be implemented in further studies; the first alternative is generating only one stream of sound, indicating only the anterior structure in order to prevent overloading the auditory processing channels. The second alternative is a selective feedback method, in which only one of the mapped data would be reported, e.g. only distance. The final variant is a step-wise feedback method, which automatically or manually moves through the mapped information. In this approach the system could first report the type of structures in front of the tool, then by choosing a focus option, more detail about the required structure, i.e. density and distance, would be mapped on the tone. This solution might contain two additional flaws; initially, the duration of the task might be further prolonged, and additionally, a highly accurate tracking and registration method is required, which brings us to the next limitation of the current setup.

The measurement of the distance and accuracy in this setup is purely done in the virtual world. Even though the 3D model is tracked and registered, the jitter and the noise in the tracking and registration system could still impose uncontrollable and unpredictable errors on our results. For further studies, it is highly encouraged that a hybrid tracking system is implemented to reduce these errors even more. This misregistration, when viewing the overlaid virtual content (e.g. virtual grasper), can frustrate the user and increase the mental process required to decode the correct position of the instrument. Even though the amount of this registration error is relatively small, due to immediate proximity of the camera to the workspace, as well as its magnifying properties, it appears more substantial in the user's eye. Alternatively, similar to the work of *Sengiku et al.*, Computer Vision can be exploited to control the correctness of the virtual kidney and 3D model registration [80]. Additionally, similar to the work of *Chen et al.* and *Simpfendörfer et al.*, external image modalities such as ultrasound or fluoroscopy can be exploited to improve the registration process [79, 82]. It should be pointed out that clamping prevents circulation in the kidney, hence the Doppler effect is mostly uninformative under these circumstances.

The generated volume for this work is a static object. Recently a new generation of runtime mesh modifiers has been introduced, which is commonly known as Procedural Voxel Engine [172].

Using this method, one can create a more realistic model of the kidney, apply deformation to the mesh, and remove unnecessary parts of the mesh at will. Additionally, combining 3D reconstruction of intraoperative images with this realtime mesh modifier could minimise the reoccurring errors in the current solutions drastically.

The accuracy of interactions between the grasper and the 3D model is yet unknown. In the follow-up studies, measuring the accuracy of the task in the real world should be considered. The registration error of the virtual and real model should be thoroughly observed and recorded as well.

A general shortcoming in medical and biomedical research, is the availability of an ideal study population, i.e. the medical experts, doctors, and surgeons. Apart from having a small population, the evaluation study of this thesis suffered from the absence of this ideal population. Having a population that could provide expert opinions in the design and implementation of the techniques can significantly impact the development process. Since the preliminary results are promising, it is now more probable to find available experts who would be willing to try and review the current setup. Acquiring this population must be considered as a major factor in the design of the follow-ups studies.

6 Summary and Conclusion

The work of this thesis, instituted to answer this research question:

“How much can auditory and visual AR display methods contribute to the amelioration of the speed, accuracy, or workload during localisation of sub-surface structures of interest in laparoscopic resection site repair phase in partial nephrectomy?”

In order to improve the current conventions of LPN/RAPN, we set out to introduce a possible AR solution for the resection site repair phase (RSRP). After providing the clinical motivation of this project, the basics of laparoscopic surgeries, visual and auditory AR displays were demonstrated. Afterwards, the current state of the art, along with the documentation of related works, was presented. Based on the provided work, an analysis of the requirement for the possible solution was carried out. Founded on related works and requirements analysis a basic concept for the possible solution to this question was formed and expanded upon.

The solution consists of two main AR approaches, i.e. visual and auditory AR displays. To provide information to the surgeon during this phase, the visual display (VD) generates an overlay of internal structures of the kidney at the tip of the surgical instrument. Alongside the VD, the auditory display (AD) provides the same information in the form of auditory feedback, indicating the type, density, and distance of the structure that is being inspected. The solution was implemented using the Unity Engine for interaction and visualisation, Pure Data for sonification, B. Braun EinsteinVision® 3.0 as laparoscopic camera, and NDI Polaris Spectra as the tracking system. A total of six displays, i.e. two visual and four auditory designs, were devised.

The proposed concepts were then evaluated, and the results were analysed in order to investigate their effects on the localisation of critical structures. For the user study, eleven medical students with the necessary knowledge of human anatomy and medical procedures were tested. This study was focused on examining the accuracy, duration, and workload rating improvements with the proposed solution, through a series of localisation tasks. During these tasks, the user was encouraged to locate the target sites via recognition of neighbouring landmarks and recognisable structures. The final results indicate that there is a significant improvement in the accuracy and workload of the procedure when the VD is enabled. Accordingly, an average improvement of 56.86% (with a mean distance of 11.131mm in the control condition and 4.801mm with the visual assistance) in the accuracy, and 23.09% (average of 14.139 for the control condition and 10.873 with the visual assistance) in the workload rating, were observed. The combination of both displays was able to achieve significant improvement of 42.21% in the localisation accuracy (with a mean distance of 6.432mm with the visio-auditory assistance). As for the AD, no significant

effect was detected, even though there were improvements in our study population. However, this does not reduce the validity of auditory AR methods, since the comparison of visual and auditory spatial models in an equal manner, requires hours of training and specified methods.

In the end, the effects of the evaluation study were discussed, and the shortcomings and possible improvements, in the context of the future work, were pointed out.

This project concludes that, as hypothesised by this work, AR methods do improve the workflow of LPN/RAPN, during the RSRP. A significant improvement in the accuracy and subjective workload was observed. Hence, introducing an AR method into the workflow of LPN/RAPN is supported by these results. Even though these significant effects were only visible in the presence of VD, AD also improved the accuracy and workload of our population, compared to the absence of AR assistance. As the preliminary prototype of a novel assistance method, the results show potential and promise to be further investigated and developed.

Bibliography

- [1] National Center for Health Statistics, & Centers for Disease Control and Prevention. Leading Causes of Death, 1900 - 1998, 1998. https://www.cdc.gov/nchs/data/statab/lead1900_98.pdf, Accessed on October 2020.
- [2] National Center for Health Statistics, & Centers for Disease Control and Prevention. Leading causes of death and numbers of deaths, by sex, race, and Hispanic origin: United States, 1980 and 2017, 2018. https://www.cdc.gov/nchs/hus/contents2018.htm#Table_006, Accessed on October 2020.
- [3] OECD/European Union 2018. Main causes of mortality, 2018. doi:https://doi.org/10.1787/health_glance_eur-2018-9-en, Accessed on October 2020.
- [4] National Cancer Institute (US). Types of Cancer Treatment, 2019. <https://www.cancer.gov/about-cancer/treatment/types>, Accessed on October 2020.
- [5] Kimberly D. Miller, Leticia Nogueira, Angela B. Mariotto, Julia H. Rowland, K. Robin Yabroff, Catherine M. Alfano, Ahmedin Jemal, Joan L. Kramer, and Rebecca L. Siegel. Cancer treatment and survivorship statistics, 2019. *CA: a cancer journal for clinicians*, 69(5):363–385, 2019. doi:10.3322/caac.21565.
- [6] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, and Ahmedin Jemal. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424, 2018. doi:10.3322/caac.21492.
- [7] World Cancer Research Fund. Kidney cancer statistics, 2018. <https://www.wcrf.org/dietandcancer/cancer-trends/kidney-cancer-statistics>, Accessed on October 2020.
- [8] National Cancer Institute (US). *PDQ Cancer Information Summaries: Renal Cell Cancer Treatment (PDQ®): Patient Version*. National Cancer Institute (US), Bethesda (MD), 2002.
- [9] Fabian Joeres, Daniel Schindele, Maria Luz, Simon Blaschke, Nele Russwinkel, Martin Schostak, and Christian Hansen. How well do software assistants for minimally invasive partial nephrectomy meet surgeon information needs? A cognitive task analysis and literature review study. *PloS one*, 14(7):e0219920, 2019. doi:10.1371/journal.pone.0219920.

- [10] Sylvain Bernhardt, Stéphane A. Nicolau, Luc Soler, and Christophe Doignon. The status of augmented reality in laparoscopic surgery as of 2016. *Medical image analysis*, 37:66–90, 2017. doi:10.1016/j.media.2017.01.007.
- [11] Dieter Schmalstieg and Tobias Höllerer. *Augmented reality: Principles and practice*. Addison-Wesley, Boston, 2016. <http://proquest.tech.safaribooksonline.de/9780133153217>.
- [12] David Black, Christian Hansen, Arya Nabavi, Ron Kikinis, and Horst Hahn. A Survey of auditory display in image-guided interventions. *International journal of computer assisted radiology and surgery*, 12(10):1665–1676, 2017. doi:10.1007/s11548-017-1547-z.
- [13] Bernhard Preim and Charl Botha. *Visual Computing for Medicine, 2nd Edition*. Morgan Kaufmann, 2nd edition edition, 2013.
- [14] Hanna Schraffenberger and Edwin van der Heide. Multimodal augmented reality - The norm rather than the exception. In Wolfgang Hürst, Daisuke Iwai, and Prabhakaran Balakrishnan, editors, *Proceedings of the 2016 workshop on Multimodal Virtual and Augmented Reality - MVAR '16*, pages 1–6, New York, New York, USA, 2016. ACM Press. doi:10.1145/3001959.3001960.
- [15] Kevin Cwach and Louis Kavoussi. Past, present, and future of laparoscopic renal surgery. *Investigative and clinical urology*, 57(Suppl 2):S110–S113, 2016. doi:10.4111/icu.2016.57.S2.S110.
- [16] Matthew T. Gettman, Michael L. Blute, George K. Chow, Richard Neururer, Georg Bartsch, and Reinhard Peschel. Robotic-assisted laparoscopic partial nephrectomy: technique and initial clinical experience with DaVinci robotic system. *Urology*, 64(5):914–918, 2004. doi:10.1016/j.urology.2004.06.049.
- [17] Jihad H. Kaouk, Ali Khalifeh, Shahab Hillyer, Georges-Pascal Haber, Robert J. Stein, and Riccardo Autorino. Robot-assisted laparoscopic partial nephrectomy: step-by-step contemporary technique and surgical outcomes at a single high-volume institution. *European urology*, 62(3):553–561, 2012. doi:10.1016/j.eururo.2012.05.021.
- [18] Christopher J. Lote. *Principles of renal physiology*. Springer, New York, 5th ed. edition, 2012.
- [19] Umberto Capitanio, Karim Bensalah, Axel Bex, Stephen A. Boorjian, Freddie Bray, Jonathan Coleman, John L. Gore, Maxine Sun, Christopher Wood, and Paul Russo. Epidemiology of Renal Cell Carcinoma. *European urology*, 75(1):74–84, 2019. doi:10.1016/j.eururo.2018.08.036.
- [20] National Cancer Institute (US). Cancer Staging, 2015. <https://www.cancer.gov/about-cancer/diagnosis-staging/staging>, Accessed on October 2020.
- [21] United Kingdom National Health Service. Laparoscopy (keyhole surgery), 2018. <https://www.nhs.uk/conditions/laparoscopy/>, Accessed on October 2020.

- [22] Udaya Kumar and Inderbir S. Gill. Learning curve in human laparoscopic surgery. *Current urology reports*, 7(2):120–124, 2006. doi:10.1007/s11934-006-0070-5.
- [23] Patricia L. Figert, Adrian E. Park, Donald B. Witzke, and Richard W. Schwartz. Transfer of training in acquiring laparoscopic skills. *Journal of the American College of Surgeons*, 193(5):533–537, 2001. doi:10.1016/s1072-7515(01)01069-9.
- [24] David C. Kerbl, Elspeth M. McDougall, Ralph V. Clayman, and Phillip Mucksavage. A history and evolution of laparoscopic nephrectomy: perspectives from the past and future directions in the surgical management of renal tumors. *The Journal of urology*, 185(3):1150–1154, 2011. doi:10.1016/j.juro.2010.10.040.
- [25] Nariman Ahmadi and Monish Aron. Laparoscopic Right Donor Nephrectomy (Transperitoneal Approach). In Mahesh Desai and Arvind P. Ganpule, editors, *Laparoscopic donor nephrectomy*, pages 59–72. Springer, Singapore, 2017. doi:10.1007/978-981-10-2849-6_{_}6.
- [26] Pascal Mouracade and Jihad Kaouk. Robot-Assisted Laparoscopic Adrenalectomy. In Hubert John and Peter Wiklund, editors, *Robotic urology*, pages 125–137. Springer, Cham, Switzerland, 2018. doi:10.1007/978-3-319-65864-3_{_}10.
- [27] Ronney Abaza. Robotic Partial Nephrectomy- Techniques for Improving Outcomes- Live Case Tips & Tricks II, 2016. <https://www.youtube.com/watch?v=Ji07K42HxF4>, Accessed on October 2020.
- [28] Ronney Abaza. Avoiding Positive Margins During Robotic Partial Nephrectomy presented by Ronney Abaza, MD - YouTube, 2018. https://www.youtube.com/watch?v=C3VTbb_1GAM, Accessed on October 2020.
- [29] James Porter. Retroperitoneal Robotic Partial Nephrectomy surgical procedure, 2015. <https://www.youtube.com/watch?v=nwrBKNbLCv8>, Accessed on October 2020.
- [30] RK Mishra. Laparoscopic Nephrectomy, 2019. <https://youtu.be/zNaBE53glfA>, Accessed on October 2020.
- [31] Adam Kibel. Robotic Assisted Laparoscopic Partial Nephrectomy, 12/11/2018. <https://www.youtube.com/watch?v=GQm90mWVMJM>, Accessed on October 2020.
- [32] Burak Turna, Rodrigo Frota, Kazumi Kamoi, Yi-Chia Lin, Monish Aron, Mihir M. Desai, Jihad H. Kaouk, and Inderbir S. Gill. Risk factor analysis of postoperative complications in laparoscopic partial nephrectomy. *The Journal of urology*, 179(4):1289–94; discussion 1294–5, 2008. doi:10.1016/j.juro.2007.11.070.
- [33] Michael A. Gorin, Rajan Ramanathan, and Raymond J. Leveillee. Laparoscopic techniques applied to open surgery: sliding-clip renorrhaphy. *Urology*, 77(3):751–753, 2011. doi:10.1016/j.j.urology.2010.10.025.

- [34] Hernan O. Altamar, Rowena E. Ong, Courtenay L. Glisson, Davis P. Viprakasit, Michael I. Miga, Stanley Duke Herrell, and Robert L. Galloway. Kidney deformation and intraprocedural registration: a study of elements of image-guided kidney surgery. *Journal of endourology*, 25(3):511–517, 2011. doi:10.1089/end.2010.0249.
- [35] Rowena E. Ong, S. Duke Herrell III, Michael I. Miga, and Robert L. Galloway Jr. A kidney deformation model for use in non-rigid registration during image-guided surgery. In *Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling*, page 69180W, San Diego, 2008. International Society for Optics and Photonics. doi:10.1117/12.771669.
- [36] I. Figueroa-Garcia, J. Peyrat, G. Hamarneh, and R. Abugharbieh. Biomechanical kidney model for predicting tumor displacement in the presence of external pressure load. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 810–813, 2014. doi:10.1109/ISBI.2014.6867994.
- [37] Steven M. Lucas, Matthew J. Mellon, Luke Erntsberger, and Chandru P. Sundaram. A comparison of robotic, laparoscopic and open partial nephrectomy. *JSLS : Journal of the Society of Laparoendoscopic Surgeons*, 16(4):581–587, 2012. doi:10.4293/108680812X13462882737177.
- [38] Hongzoo Park, Seok-Soo Byun, Hyeon Hoe Kim, Seung Bae Lee, Tae Gyun Kwon, Seung Hyun Jeon, Seok Ho Kang, Seong Il Seo, Tae Hee Oh, Youn Soo Jeon, Wan Lee, Tae-Kon Hwang, Koon Ho Rha, Ill Young Seo, Dong Deuk Kwon, Yong-June Kim, Yunhee Choi, and Sue Kyung Park. Comparison of laparoscopic and open partial nephrectomies in t1a renal cell carcinoma: a korean multicenter experience. *Korean journal of urology*, 51(7):467–471, 2010. doi:10.4111/kju.2010.51.7.467.
- [39] Inderbir S. Gill, Mihir M. Desai, Jihad H. Kaouk, Anoop M. Meraney, David P. Murphy, Gyung Tak Sung, and Andrew C. Novick. Laparoscopic partial nephrectomy for renal tumor: duplicating open surgical techniques. *The Journal of urology*, 167(2 Pt 1):469–7; discussion 475–6, 2002.
- [40] Hyeon Jun Jang, Wan Song, Yoon Seok Suh, U. Seok Jeong, Hwang Gyun Jeon, Byong Chang Jeong, Seong Soo Jeon, Hyun Moo Lee, Han Yong Choi, and Seong Il Seo. Comparison of perioperative outcomes of robotic versus laparoscopic partial nephrectomy for complex renal tumors (RENAL nephrometry score of 7 or higher). *Korean journal of urology*, 55(12):808–813, 2014.
- [41] Xiaolong Zhang, Jiajun Yan, Yu Ren, Chong Shen, Xiangrong Ying, and Shouhua Pan. Robot-assisted versus laparoscopic partial nephrectomy for localized renal tumors: a meta-analysis. *International Journal of Clinical and Experimental Medicine*, 7(12):4770–4779, 2014.
- [42] Ji Eun Choi, Ji Hye You, Dae Keun Kim, Koon Ho Rha, and Seon Heui Lee. Comparison of perioperative outcomes between robotic and laparoscopic partial nephrectomy: a systematic

-
- review and meta-analysis. *European urology*, 67(5):891–901, 2015. doi:10.1016/j.eururo.2014.12.028.
- [43] R. Berguer, W. D. Smith, and Y. H. Chung. Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery. *Surgical endoscopy*, 15(10):1204–1207, 2001. doi:10.1007/s004640080030.
- [44] Sonal Arora, Nick Sevdalis, Debra Nestel, Maria Woloshynowych, Ara Darzi, and Roger Kneebone. The impact of stress on surgical performance: a systematic review of the literature. *Surgery*, 147(3):318–30, 330.e1–6, 2010. doi:10.1016/j.surg.2009.10.007.
- [45] Adrian Park, Gyusung Lee, F. Jacob Seagull, Nora Meenaghan, and David Dexter. Patients benefit while surgeons suffer: an impending epidemic. *Journal of the American College of Surgeons*, 210(3):306–313, 2010. doi:10.1016/j.jamcollsurg.2009.10.017.
- [46] Kim Platte, Chantal C. J. Alleblas, Joanna Inthout, and Theodoor E. Nieboer. Measuring fatigue and stress in laparoscopic surgery: validity and reliability of the star-track test. *Minimally invasive therapy & allied technologies : MITAT : official journal of the Society for Minimally Invasive Therapy*, 28(1):57–64, 2019. doi:10.1080/13645706.2018.1470984.
- [47] Harsimrat Singh, Hemel N. Modi, Samridha Ranjan, James W. R. Dille, Dimitrios Airantzis, Guang-Zhong Yang, Ara Darzi, and Daniel R. Leff. Robotic Surgery Improves Technical Performance and Enhances Prefrontal Activation During High Temporal Demand. *Annals of Biomedical Engineering*, 46(10):1621–1636, 2018. doi:10.1007/s10439-018-2049-z.
- [48] Esther Lau, Nawar A. Alkhamesi, and Christopher M. Schlachta. Impact of robotic assistance on mental workload and cognitive performance of surgical trainees performing a complex minimally invasive suturing task. *Surgical endoscopy*, 34(6):2551–2559, 2020. doi:10.1007/s00464-019-07038-9.
- [49] P. Milgram and F. Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems*, 77:1321–1329, 1994.
- [50] Ronald T. Azuma. A Survey of Augmented Reality: Presence: Teleoperators and Virtual Environments, 6(4), 355–385. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997. doi:10.1162/PRES.1997.6.4.355.
- [51] MJ Anderson. Augmented or Virtual: How Do You Like Your Reality?, 2015. <https://www.trekk.com/insights/augmented-or-virtual-how-do-you-your-reality>, Accessed on October 2020.
- [52] Art - Advance Realtime Tracking. Markers - Markers & Targets - Products - ART Advanced Realtime Tracking, 10/16/2020. <https://ar-tracking.com/products/markers-targets/markers/>, Accessed on October 2020.
-

- [53] Oliver Bimber and Ramesh Raskar. *Modern Approaches to Augmented Reality*. ACM SIGGRAPH Courses. Association for Computing Machinery, New York, NY, United States, 2006. doi:10.1145/1185657.1185796.
- [54] D.W.F. van Krevelen and R. Poelman. A Survey of Augmented Reality Technologies, Applications and Limitations: International Journal of Virtual Reality, 9(2), 1-20. *International Journal of Virtual Reality*, 9(2):1–20, 2019. doi:10.20870/IJVR.2010.9.2.2767.
- [55] Maki Sugimoto, Hideki Yasuda, Keiji Koda, Masato Suzuki, Masato Yamazaki, Tohru Tezuka, Chihiro Kosugi, Ryota Higuchi, Yoshihisa Watayo, Yohsuke Yagawa, Shuichiro Uemura, Hironori Tsuchiya, and Takeshi Azuma. Image overlay navigation by markerless surface registration in gastrointestinal, hepatobiliary and pancreatic surgery. *Journal of hepatobiliary-pancreatic sciences*, 17(5):629–636, 2010. doi:10.1007/s00534-009-0199-y.
- [56] Benjamin J. Dixon, Michael J. Daly, Harley Chan, Allan D. Vescan, Ian J. Witterick, and Jonathan C. Irish. Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. *Surgical endoscopy*, 27(2):454–461, 2013. doi:10.1007/s00464-012-2457-3.
- [57] Hasaneen Fathy Al Janabi, Abdullatif Aydin, Sharanya Palaneer, Nicola Macchione, Ahmed Al-Jabir, Muhammad Shamim Khan, Prokar Dasgupta, and Kamran Ahmed. Effectiveness of the HoloLens mixed-reality headset in minimally invasive surgery: a simulation-based feasibility study. *Surgical Endoscopy*, 34(3):1143–1149, 2020. doi:10.1007/s00464-019-06862-3.
- [58] Stephen Barrass. *Auditory Information Design*. PhD Thesis, The Australian National University, 1997. <https://dl.acm.org/doi/book/10.5555/929501>.
- [59] G. Kramer, Bruce Walker, T. Bonebright, P. Cook, and S. Tipei. *The sonification report: Status of the field and research agenda*. International Community for Auditory Display, 1999.
- [60] Thomas Hermann, Andy Hunt, and John G. Neuhoff. *The sonification handbook*. Logos Verlag, Berlin, 2011.
- [61] Gaetan Parseihian, Charles Gondre, Mitsuko Aramaki, Solvi Ystad, and Richard Kronland-Martinet. Comparison and Evaluation of Sonification Strategies for Guidance Tasks. *IEEE Transactions on Multimedia*, 18(4):674–686, 2016. doi:10.1109/TMM.2016.2531978.
- [62] Osamu Ukimura, Masahiko Nakamoto, and Inderbir S. Gill. Three-dimensional reconstruction of renovascular-tumor anatomy to facilitate zero-ischemia partial nephrectomy. *European urology*, 61(1):211–217, 2012. doi:10.1016/j.eururo.2011.07.068.
- [63] Amrith Raj Rao, Robert Gray, Erik Mayer, Hanif Motiwala, Marc Laniado, and Omer Karim. Occlusion angiography using intraoperative contrast-enhanced ultrasound scan (CEUS): a novel technique demonstrating segmental renal blood supply to assist zero-ischaemia robot-assisted partial nephrectomy. *European urology*, 63(5):913–919, 2013. doi:10.1016/j.eururo.2012.10.034.

- [64] Junya Furukawa, Hideaki Miyake, Kazushi Tanaka, Maki Sugimoto, and Masato Fujisawa. Console-integrated real-time three-dimensional image overlay navigation for robot-assisted partial nephrectomy with selective arterial clamping: early single-centre experience with 17 cases. *The international journal of medical robotics + computer assisted surgery : MRCAS*, 10(4):385–390, 2014. doi:10.1002/rcs.1574.
- [65] Satoshi Kobayashi, Byunghyun Cho, Arnaud Huaultmé, Katsunori Tatsugami, Hiroshi Honda, Pierre Jannin, Makoto Hashizumea, and Masatoshi Eto. Assessment of surgical skills by using surgical navigation in robot-assisted partial nephrectomy. *International journal of computer assisted radiology and surgery*, 14(8):1449–1459, 2019. doi:10.1007/s11548-019-01980-8.
- [66] Francesco Porpiglia, Cristian Fiori, Enrico Checcucci, Daniele Amparore, and Riccardo Bertolo. Hyperaccuracy Three-dimensional Reconstruction Is Able to Maximize the Efficacy of Selective Clamping During Robot-assisted Partial Nephrectomy for Complex Renal Masses. *European urology*, 74(5):651–660, 2018. doi:10.1016/j.eururo.2017.12.027.
- [67] Kazuhiro Nakamura, Yukio Naya, Satoki Zenbutsu, Kazuhiro Araki, Shuko Cho, Sho Ohta, Naoki Nihei, Hiroyoshi Suzuki, Tomohiko Ichikawa, and Tatsuo Igarashi. Surgical navigation using three-dimensional computed tomography images fused intraoperatively with live video. *Journal of endourology*, 24(4):521–524, 2010. doi:10.1089/end.2009.0365.
- [68] Dongwen Wang, Bin Zhang, Xiaobin Yuan, Xuhui Zhang, and Chen Liu. Preoperative planning and real-time assisted navigation by three-dimensional individual digital model in partial nephrectomy with three-dimensional laparoscopic system. *International journal of computer assisted radiology and surgery*, 10(9):1461–1468, 2015. doi:10.1007/s11548-015-1148-7.
- [69] Elias S. Hyams, Mark Perlmutter, and Michael D. Stifelman. A prospective evaluation of the utility of laparoscopic Doppler technology during minimally invasive partial nephrectomy. *Urology*, 77(3):617–620, 2011. doi:10.1016/j.urology.2010.05.011.
- [70] Scott Tobis, Joy Knopf, Christopher Silvers, Jorge Yao, Hani Rashid, Guan Wu, and Dragan Golijanin. Near infrared fluorescence imaging with robotic assisted laparoscopic partial nephrectomy: initial clinical experience for renal cortical tumors. *The Journal of urology*, 186(1):47–52, 2011. doi:10.1016/j.juro.2011.02.2701.
- [71] Alborz Amir-Khalili, Ghassan Hamarneh, Jean-Marc Peyrat, Julien Abinahed, Osama Al-Alao, Abdulla Al-Ansari, and Rafeef Abugharbieh. Automatic segmentation of occluded vasculature via pulsatile motion analysis in endoscopic robot-assisted partial nephrectomy video. *Medical image analysis*, 25(1):103–110, 2015. doi:10.1016/j.media.2015.04.010.
- [72] Adam C. Mues, Zhamshid Okhunov, Ketan Badani, Mantu Gupta, and Jaime Landman. Intraoperative evaluation of renal blood flow during laparoscopic partial nephrectomy with a novel Doppler system. *Journal of endourology*, 24(12):1953–1956, 2010. doi:10.1089/end.2010.0171.

- [73] Ahmad Alenezi, Aamir Motiwala, Susannah Eves, Rob Gray, Asha Thomas, Isabelle Meiers, Haytham Sharif, Hanif Motiwala, Marc Laniado, and Omer Karim. Robotic assisted laparoscopic partial nephrectomy using contrast-enhanced ultrasound scan to map renal blood flow. *The international journal of medical robotics + computer assisted surgery : MRCAS*, 13(1), 2017. doi:10.1002/rcs.1738.
- [74] Michael S. Borofsky, Inderbir S. Gill, Ashok K. Hemal, Tracy P. Marien, Isuru Jayaratna, Louis S. Krane, and Michael D. Stifelman. Near-infrared fluorescence imaging to facilitate super-selective arterial clamping during zero-ischæmia robotic partial nephrectomy. *BJU international*, 111(4):604–610, 2013. doi:10.1111/j.1464-410X.2012.11490.x.
- [75] Marc A. Bjurlin, Melanie Gan, Tyler R. McClintock, Alessandro Volpe, Michael S. Borofsky, Alexandre Mottrie, and Michael D. Stifelman. Near-infrared fluorescence imaging: emerging applications in robotic upper urinary tract surgery. *European urology*, 65(4):793–801, 2014. doi:10.1016/j.eururo.2013.09.023.
- [76] Archie Hughes-Hallett, Philip Pratt, Erik Mayer, Shirley Martin, Ara Darzi, and Justin Vale. Image guidance for all–TilePro display of 3-dimensionally reconstructed images in robotic partial nephrectomy. *Urology*, 84(1):237–242, 2014. doi:10.1016/j.urology.2014.02.051.
- [77] Fubo Wang, Chao Zhang, Fei Guo, Jin Ji, Ji Lyu, Zhi Cao, and Bo Yang. Navigation of Intelligent/Interactive Qualitative and Quantitative Analysis Three-Dimensional Reconstruction Technique in Laparoscopic or Robotic Assisted Partial Nephrectomy for Renal Hilar Tumors. *Journal of endourology*, 33(8):641–646, 2019. doi:10.1089/end.2018.0570.
- [78] Philip Pratt, Erik Mayer, Justin Vale, Daniel Cohen, Eddie Edwards, Ara Darzi, and Guang-Zhong Yang. An effective visualisation and registration system for image-guided robotic partial nephrectomy. *Journal of robotic surgery*, 6(1):23–31, 2012. doi:10.1007/s11701-011-0334-z.
- [79] Yuanbo Chen, Hulin Li, Dingtao Wu, Keming Bi, and Chunxiao Liu. Surgical planning and manual image fusion based on 3D model facilitate laparoscopic partial nephrectomy for intrarenal tumors. *World journal of urology*, 32(6):1493–1499, 2014. doi:10.1007/s00345-013-1222-0.
- [80] Atsushi Sengiku, Masanao Koeda, Atsuro Sawada, Jin Kono, Naoki Terada, Toshinari Yamasaki, Kiminori Mizushino, Takahiro Kunii, Katsuhiko Onishi, Hiroshi Noborio, and Osamu Ogawa. Augmented Reality Navigation System for Robot-Assisted Laparoscopic Partial Nephrectomy. In Aaron Marcus and Wentao Wang, editors, *Design, user experience, and usability*, LNCS Sublibrary: SL3 - Information Systems and Applications, incl. Internet/Web, and HCI, pages 575–584, Cham, Switzerland, 2017. Springer. https://doi.org/10.1007/978-3-319-58637-3_45.
- [81] Dogu Teber, Selcuk Guven, Tobias Simpfendorfer, Mathias Baumhauer, Esref Oguz Guven, Faruk Yencilek, Ali Serdar Gozen, and Jens Rassweiler. Augmented reality: a new tool to

- improve surgical accuracy during laparoscopic partial nephrectomy? Preliminary in vitro and in vivo results. *European urology*, 56(2):332–338, 2009. doi:10.1016/j.eururo.2009.05.017.
- [82] Tobias Simpfendorfer, Claudia Gasch, Gencay Hatiboglu, Michael Müller, Lena Maier-Hein, Markus Hohenfellner, and Dogu Teber. Intraoperative Computed Tomography Imaging for Navigated Laparoscopic Renal Surgery: First Clinical Experience. *Journal of endourology*, 30(10):1105–1111, 2016. doi:10.1089/end.2016.0385.
- [83] Dexin Dong, Zhigang Ji, Hanzhong Li, Weigang Yan, and Yushi Zhang. Laparoscopic Nephron Sparing Surgery Assisted with Laparoscopic Ultrasonography on Centrally Located Renal Tumor - Single Center Experience. *Urologia internationalis*, 97(2):195–199, 2016. doi:10.1159/000446026.
- [84] Roberta Gunelli, Massimo Fiori, Cristiano Salaris, Umberto Salomone, Marco Urbinati, Alexia Vici, Teo Zenico, and Mauro Bertocco. The role of intraoperative ultrasound in small renal mass robotic enucleation. *Archivio italiano di urologia, andrologia : organo ufficiale [di] Societa italiana di ecografia urologica e nefrologica*, 88(4):311–313, 2016. doi:10.4081/aiua.2016.4.311.
- [85] James Jeffery Reeves, Andrew Forauer, John D. Seigne, and Elias S. Hyams. Image-Guided Embolization Coil Placement for Identification of an Endophytic, Isoechoic Renal Mass During Robotic Partial Nephrectomy. *Journal of endourology case reports*, 1(1):59–61, 2015. doi:10.1089/cren.2015.0022.
- [86] Carling L. Cheung, Chris Wedlake, John Moore, Stephen E. Pautler, and Terry M. Peters. Fused video and ultrasound images for minimally invasive partial nephrectomy: a phantom study. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 13(Pt 3):408–415, 2010. doi:10.1007/978-3-642-15711-0_{ }51.
- [87] Philip Pratt, Alexander Jaeger, Archie Hughes-Hallett, Erik Mayer, Justin Vale, Ara Darzi, Terry Peters, and Guang-Zhong Yang. Robust ultrasound probe tracking: initial clinical experiences during robot-assisted partial nephrectomy. *International journal of computer assisted radiology and surgery*, 10(12):1905–1913, 2015. doi:10.1007/s11548-015-1279-x.
- [88] M. Raschid Hoda and Galf Popken. Surgical outcomes of fluorescence-guided laparoscopic partial nephrectomy using 5-aminolevulinic acid-induced protoporphyrin IX. *The Journal of surgical research*, 154(2):220–225, 2009. doi:10.1016/j.jss.2008.12.027.
- [89] Jeremy Kawahara, Jean-Marc Peyrat, Julien Abinahed, Osama Al-Alao, Abdulla Al-Ansari, Rafeef Abugharbieh, and Ghassan Hamarneh. Automatic labelling of tumourous frames in free-hand laparoscopic ultrasound video. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 17(Pt 2):676–683, 2014. doi:10.1007/978-3-319-10470-6_{ }84.

- [90] Osamu Ukimura and Inderbir S. Gill. Imaging-assisted endoscopic surgery: Cleveland Clinic experience. *Journal of endourology*, 22(4):803–810, 2008. doi:10.1089/end.2007.9823.
- [91] Pauline Chauvet, Toby Collins, Clement Debize, Lorraine Novais-Gameiro, Bruno Pereira, Adrien Bartoli, Michel Canis, and Nicolas Bourdel. Augmented reality in a tumor resection model. *Surgical endoscopy*, 32(3):1192–1201, 2018. doi:10.1007/s00464-017-5791-7.
- [92] Philip Edgcumbe, Rohit Singla, Philip Pratt, Caitlin Schneider, Christopher Nguan, and Robert Rohling. Follow the light: projector-based augmented reality intracorporeal system for laparoscopic surgery. *Journal of medical imaging (Bellingham, Wash.)*, 5(2):021216, 2018. doi:10.1117/1.JMI.5.2.021216.
- [93] Alborz Amir-Khalili, Masoud S. Nosrati, Jean-Marc Peyrat, Ghassan Hamarneh, and Rafeef Abugharbieh. Uncertainty-Encoded Augmented Reality for Robot-Assisted Partial Nephrectomy: A Phantom Study. In Hongen Liao, editor, *Augmented reality environments for medical imaging and computer-assisted interventions*, LNCS sublibrary. SL 6 - Image processing, computer vision, pattern recognition, and graphics, pages 182–191, Heidelberg, 2013. Springer.
- [94] Rohit Singla, Philip Edgcumbe, Philip Pratt, Christopher Nguan, and Robert Rohling. Intra-operative ultrasound-based augmented reality guidance for laparoscopic surgery. *Healthcare technology letters*, 4(5):204–209, 2017. doi:10.1049/htl.2017.0063.
- [95] Arnaud Doerfler, Abeni Oitchayomi, and Xavier Tillou. A simple method for ensuring resection margins during laparoscopic partial nephrectomy: the intracorporeal ultrasonography. *Urology*, 84(5):1240–1242, 2014. doi:10.1016/j.urology.2014.07.025.
- [96] Marta Kersten-Oertel, Pierre Jannin, and D. Louis Collins. The state of the art of visualization in mixed reality image guided surgery. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, 37(2):98–112, 2013. doi:10.1016/j.compmedimag.2013.01.009.
- [97] Rositsa Bogdanova, Pierre Boulanger, and Bin Zheng. Depth Perception of Surgeons in Minimally Invasive Surgery. *Surgical innovation*, 23(5):515–524, 2016. doi:10.1177/1553350616639141.
- [98] Felix Bork, Bernhard Fuers, Anja-Katharina Schneider, Francisco Pinto, Christoph Graumann, and Nassir Navab. Auditory and Visio-Temporal Distance Coding for 3-Dimensional Perception in Medical Augmented Reality. In *2015 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 7–12. IEEE / Institute of Electrical and Electronics Engineers Incorporated, 2015. doi:10.1109/ISMAR.2015.16.
- [99] Joseph Plazak, Daniel A. DiGiovanni, D. Louis Collins, and Marta Kersten-Oertel. Cognitive load associations when utilizing auditory display within image-guided neurosurgery. *International journal of computer assisted radiology and surgery*, 14(8):1431–1438, 2019. doi:10.1007/s11548-019-01970-w.

-
- [100] Hyunseok Choi, Byunghyun Cho, Ken Masamune, Makoto Hashizume, and Jaesung Hong. An effective visualization technique for depth perception in augmented reality-based surgical navigation. *The international journal of medical robotics + computer assisted surgery : MRCAS*, 12(1):62–72, 2016. doi:10.1002/rcs.1657.
- [101] Timo Ropinski, Frank Steinicke, and Klaus Hinrichs. Visually Supporting Depth Perception in Angiography Imaging. In Andreas Butz, Brian Fisher, Antonio Krüger, and Patrick Olivier, editors, *Smart Graphics (vol. [4073])*, Lecture Notes in Computer Science, pages 93–104, Berlin Heidelberg, 2006. Springer-Verlag.
- [102] Alark Joshi, Xiaoning Qian, Donald P. Dione, Ketan R. Bulsara, Christopher K. Breuer, Albert J. Sinusas, and Xenophon Papademetris. Effective visualization of complex vascular structures using a non-parametric vessel detection method. *IEEE transactions on visualization and computer graphics*, 14(6):1603–1610, 2008. doi:10.1109/TVCG.2008.123.
- [103] Marta Kersten-Oertel, Sean Jy-Shyang Chen, and D. Louis Collins. An evaluation of depth enhancing perceptual cues for vascular volume visualization in neurosurgery. *IEEE transactions on visualization and computer graphics*, 20(3):391–403, 2014. doi:10.1109/TVCG.2013.240.
- [104] Christoph Bichlmeier, Tobias Sielhorst, Sandro M. Heining, and Nassir Navab. Improving Depth Perception in Medical AR. In Alexander Horsch, editor, *Bildverarbeitung für die Medizin 2007*, Informatik aktuell, pages 217–221. Springer, Berlin, 2007. doi:10.1007/978-3-540-71091-2_{_}44.
- [105] Richard A. Steenblik. The Chromostereoscopic Process: A Novel Single Image Stereoscopic Process. doi:10.1117/12.940117.
- [106] Kai Lawonn, Maria Luz, Bernhard Preim, and Christian Hansen. Illustrative Visualization of Vascular Models for Static 2D Representations. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical image computing and computer-assisted intervention - MICCAI 2015*, LNCS sublibrary: SL6 - Image processing, computer vision, pattern recognition, and graphics, pages 399–406, Cham, 2015. Springer.
- [107] Gerd Schmidt. *Augmented Reality-Konzepte zur Verbesserung der Distanzeinschaetzung fuer die Navigation von chirurgischen Instrumenten*. Master’s Thesis, Otto-von-Guericke-Universität, Magdeburg, March, 2019. http://www.var.ovgu.de/pub/2019_MA_Schmidt_geschwaerzt.pdf.
- [108] Masaya Kitagawa, Daniell Dokko, Allison M. Okamura, and David D. Yuh. Effect of sensory substitution on suture-manipulation forces for robotic surgical systems. *The Journal of thoracic and cardiovascular surgery*, 129(1):151–158, 2005. doi:10.1016/j.jtcvs.2004.05.029.
- [109] P. W. A. Willems, H. J. Noordmans, J. J. van Overbeeke, M. A. Viergever, C. A. F. Tulleken, and J. W. Berkelbach van der Sprenkel. The impact of auditory feedback
-

- on neuronavigation. *Acta neurochirurgica*, 147(2):167–73; discussion 173, 2005. doi:10.1007/s00701-004-0412-3.
- [110] Peter A. Woerdeman, Peter W. A. Willems, Herke Jan Noordmans, and Jan Willem Berkelbach van der Sprenkel. Auditory feedback during frameless image-guided surgery in a phantom model and initial clinical experience. *Journal of neurosurgery*, 110(2):257–262, 2009. doi:10.3171/2008.3.17431.
- [111] G. Strauss, S. Schaller, B. Zamminer, S. Heininger, M. Hofer, D. Manzey, J. Meixensberger, A. Dietz, and T. C. Lüth. Klinische Erfahrungen mit einem Kollisionswarnsystem: Instrumentennavigation in der endo- und transnasalen Chirurgie. *HNO*, 59(5):470–479, 2011. doi:10.1007/s00106-010-2237-0.
- [112] Eduard H. J. Voormolen, Peter A. Woerdeman, Marijn van Stralen, Herke Jan Noordmans, Max A. Viergever, Luca Regli, and Jan Willem Berkelbach van der Sprenkel. Validation of exposure visualization and audible distance emission for navigated temporal bone drilling in phantoms. *PLOS ONE*, 7(7):e41262, 2012. doi:10.1371/journal.pone.0041262.
- [113] Byunghyun Cho, Masamichi Oka, Nozomu Matsumoto, Riichi Ouchida, Jaesung Hong, and Makoto Hashizume. Warning navigation system using real-time safe region monitoring for otologic surgery. *International journal of computer assisted radiology and surgery*, 8(3):395–405, 2013. doi:10.1007/s11548-012-0797-z.
- [114] Byunghyun Cho, Nozomu Matsumoto, Shizuo Komune, and Makoto Hashizume. A surgical navigation system for guiding exact cochleostomy using auditory feedback: a clinical feasibility study. *BioMed research international*, 2014:769659, 2014. doi:10.1155/2014/769659.
- [115] Benjamin J. Dixon, Michael J. Daly, Harley Chan, Allan Vescan, Ian J. Witterick, and Jonathan C. Irish. Augmented real-time navigation with critical structure proximity alerts for endoscopic skull base surgery. *The Laryngoscope*, 124(4):853–859, 2014. doi:10.1002/lary.24385.
- [116] C. M. Wegner and D. B. Karron. Surgical navigation using audio feedback. *Studies in health technology and informatics*, 39:450–458, 1997.
- [117] Kristen Wegner. Surgical Navigation System and Method Using Audio Feedback. In *International Conference on Auditory Display, 1998*. BCS Learning & Development Ltd., Swindon, United Kingdom, 1998. doi:10.14236/ewic/AD1998.31.
- [118] Christian Hansen, David Black, Christoph Lange, Fabian Rieber, Wolfram Lamadé, Marcello Donati, Karl J. Oldhafer, and Horst K. Hahn. Auditory support for resection guidance in navigated liver surgery. *The international journal of medical robotics + computer assisted surgery : MRCAS*, 9(1):36–43, 2013. doi:10.1002/racs.1466.
- [119] David Black, Jumana Al Issawi, Christian Hansen, Christian Rieder, and H. Hahn. Auditory support for navigated radiofrequency ablation. In *CURAC*, 2013.

- [120] Jonathan D. Katz. Noise in the operating room. *Anesthesiology*, 121(4):894–898, 2014. doi:10.1097/ALN.0000000000000319.
- [121] Masoumeh Dorri Giv, Karim Ghazikhanlou Sani, Majid Alizadeh, Ali Valinejadi, and Hesamedin Askari Majdabadi. Evaluation of noise pollution level in the operating rooms of hospitals: A study in Iran. *Interventional medicine & applied science*, 9(2):61–66, 2017. doi:10.1556/1646.9.2017.2.15.
- [122] David Black, Julian Hettig, Maria Luz, Christian Hansen, Ron Kikinis, and Horst Hahn. Auditory feedback to support image-guided medical needle placement. *International journal of computer assisted radiology and surgery*, 12(9):1655–1663, 2017. doi:10.1007/s11548-017-1537-1.
- [123] Florian Heinrich. *Augmented-Reality-Visualisierung einer Nadelnavigation im Magnetresonanztomographen*. Master’s Thesis, Otto-von-Guericke-Universität, Magdeburg, April, 2017.
- [124] International Organization for Standardization. Ergonomics of human-system interaction: Part 11: Usability: Definitions and concepts, March 2018. <https://www.iso.org/standard/63500.html>.
- [125] Paul Read and Nicola Mazzanti. *Restoration of motion picture film*. Butterworth-Heinemann series in conservation and museology. Butterworth-Heinemann, Oxford, 1. Aufl. edition, 1998. <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10175505>.
- [126] Amin Banitalebi-Dehkordi, Mahsa T. Pourazad, and Panos Nasiopoulos. The Effect of Frame Rate on 3D Video Quality and Bitrate. *3D Research*, 6(1):1–13, 2015. doi:10.1007/s13319-014-0034-3.
- [127] Tonia Mielke. *Entwicklung eines Registrierungskonzepts für laparoskopische Augmented Reality*. Bachelor’s Thesis, Otto-von-Guericke-Universität Magdeburg, Magdeburg, September, 2020.
- [128] Yolanda Vazquez-Alvarez, Ian Oakley, and Stephen A. Brewster. Auditory display design for exploration in mobile audio-augmented reality. *Personal and Ubiquitous Computing*, 16(8):987–999, 2012. doi:10.1007/s00779-011-0459-0.
- [129] vhv.rs. Orthographic Projection Camera, 2016. https://www.vhv.rs/viewpic/iwhJhTo_orthographic-projection-camera-hd-png-download/, Accessed on October 2020.
- [130] Jan Kremers, Rigmor C. Baraas, and N. Justin Marshall. *Human Color Vision*. Springer International Publishing, Cham, 2016. doi:10.1007/978-3-319-44978-4.
- [131] A. Ajmal, C. Hollitt, M. Frean, and H. Al-Sahaf. A Comparison of RGB and HSV Colour Spaces for Visual Attention Models. In *2018 International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pages 1–6, 2018. doi:10.1109/IVCNZ.2018.8634752.

- [132] Kenneth R. Alexander and Michael S. Shansky. Influence of hue, value, and chroma on the perceived heaviness of colors. *Perception & Psychophysics*, 19(1):72–74, 1976. doi:10.3758/BF03199388.
- [133] A. Vadivel, Shamik Sural, and Arun K. Majumdar. Human color perception in the HSV space and its application in histogram generation for image retrieval. In Reiner Eschbach and Gabriel G. Marcu, editors, *Color Imaging X: Processing, Hardcopy, and Applications*, SPIE Proceedings, page 598. SPIE, 2005. doi:10.1117/12.586823.
- [134] Ofir Pele and Michael Werman. The Quadratic-Chi Histogram Distance Family. In Kostas Daniilidis, Petros Maragos, and Nikos Paragios, editors, *Computer vision - ECCV 2010*, volume 6312 of *LNCS sublibrary: SL 6 - Image processing, computer vision, pattern recognition, and graphics*, pages 749–762. Springer, Berlin, 2010. doi:10.1007/978-3-642-15552-9_{ }54.
- [135] Pedro A. Marín-Reyes, Javier Lorenzo-Navarro, and Modesto Castrillón-Santana. Comparative study of histogram distance measures for re-identification. <http://arxiv.org/pdf/1611.08134v1>.
- [136] MathWorks. Matlab. <https://www.mathworks.com/products/matlab.html>.
- [137] Karen Allen and Jim Blascovich. Effects of Music on Cardiovascular Reactivity Among Surgeons. *JAMA: The Journal of the American Medical Association*, 272(11):882, 1994. doi:10.1001/jama.1994.03520110062030.
- [138] Northern Digital Inc. Polaris Spectra and Vicra, 3/13/2018. <https://www.ndigital.com/medical/products/polaris-family/>, Accessed on October 2020.
- [139] B. Braun AG. EinsteinVision® 3.0 - 3D camera system in laparoscopic and cardio-thoracic surgery. <https://www.bbraun.com/en/products/b/einsteinvision.html>, Accessed on October 2020.
- [140] M. Puckette. Pure Data: another integrated computer music environment, 1997. <https://puredata.info/docs/articles/puredata1997>, Accessed on October 2020.
- [141] Unity Technologies. Unity Engine, 2007 - 2020. <https://unity.com/>, Accessed on October 2020.
- [142] Niall Moody. LibPdIntegration, 2018. <https://github.com/LibPdIntegration>, Accessed on October 2020.
- [143] Nicholas Heller, Niranjana Sathianathan, Arveen Kalapara, Edward Walczak, Keenan Moore, Heather Kaluzniak, Joel Rosenberg, Paul Blake, Zachary Rengel, Makinna Oestreich, Joshua Dean, Michael Tradewell, Aneri Shah, Resha Tejpaul, Zachary Edgerton, Matthew Peterson, Shaneabbas Raza, Subodh Regmi, Nikolaos Papanikolopoulos, and Christopher Weight. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes. <http://arxiv.org/pdf/1904.00445v2>.

- [144] MeVis Medical Solutions AG. MeVisLab, 2020. <https://www.mevislab.de/>.
- [145] BWH and 3D Slicer contributors. 3D Slicer, 8/7/2020. <https://www.slicer.org/>, Accessed on October 2020.
- [146] Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, John Buatti, Stephen Aylward, James V. Miller, Steve Pieper, and Ron Kikinis. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging*, 30(9):1323–1341, 2012. doi:10.1016/j.mri.2012.05.001.
- [147] community Blender Foundation. Blender, 2020. <https://www.blender.org/>.
- [148] Stephen M. Casner and B. Gore. Measuring and Evaluating Workload: A Primer. In *Task Book: Biological & Physical Sciences Division and Human Research Program*. NASA/TM, Washington D.C, 2010. https://matb-files.larc.nasa.gov/Workload_Primer_TM_Final.pdf.
- [149] Aidan Byrne. Measurement of mental workload in clinical medicine: a review study. *Anesthesiology and pain medicine*, 1(2):90–94, 2011. doi:10.5812/kowsar.22287523.2045.
- [150] Susana Rubio, Eva Diaz, Jesus Martin, and Jose M. Puente. Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, 53(1):61–86, 2004. doi:10.1111/j.1464-0597.2004.00161.x.
- [151] Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Peter A. Hancock and Najmedin Meshkati, editors, *Advances in Psychology : Human Mental Workload*, volume 52, pages 139–183. North-Holland, 1988. doi:10.1016/S0166-4115(08)62386-9.
- [152] Andy P. Field, Jeremy Miles, and Zoë Field. *Discovering statistics using R*. SAGE, London, 2012.
- [153] Andy P. Field. *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll / Andy Field*. SAGE, London, 4th ed. edition, 2013.
- [154] Charles J. Spence and Jon Driver. Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, 20(3):555–574, 1994. doi:10.1037/0096-1523.20.3.555.
- [155] Jon Driver and Charles Spence. Crossmodal attention. *Current Opinion in Neurobiology*, 8(2):245–253, 1998. doi:10.1016/S0959-4388(98)80147-5.
- [156] Brenda Rapp and Sharma K. Hendel. Principles of cross-modal competition: evidence from deficits of attention. *Psychonomic Bulletin & Review*, 10(1):210–219, 2003. doi:10.3758/BF03196487.

- [157] Thomas M. van Vleet and Lynn C. Robertson. Cross-modal interactions in time and space: auditory influence on visual attention in hemispatial neglect. *Journal of cognitive neuroscience*, 18(8):1368–1379, 2006. doi:10.1162/jocn.2006.18.8.1368.
- [158] David J. Prime, John J. McDonald, Jessica Green, and Lawrence M. Ward. When cross-modal spatial attention fails. *Canadian journal of experimental psychology = Revue canadienne de psychologie experimentale*, 62(3):192–197, 2008. doi:10.1037/1196-1961.62.3.192.
- [159] David H. Warren. Early vs. Late Vision: The Role of Early Vision in Spatial Reference Systems. *Journal of Visual Impairment & Blindness*, 68(4):157–162, 1974. doi:10.1177/0145482X7406800404.
- [160] HARRY MCGURK and JOHN MACDONALD. Hearing lips and seeing voices. *Nature*, 264(5588):746–748, 1976. doi:10.1038/264746a0.
- [161] Frank Bremmer, François Klam, Jean-René Duhamel, Suliann Ben Hamed, and Werner Graf. Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *The European journal of neuroscience*, 16(8):1569–1586, 2002. doi:10.1046/j.1460-9568.2002.02206.x.
- [162] Ilana B. Witten and Eric I. Knudsen. Why Seeing Is Believing: Merging Auditory and Visual Worlds. *Neuron*, 48(3):489–496, 2005. doi:10.1016/j.neuron.2005.10.020.
- [163] Sharon E. Guttman, Lee A. Gilroy, and Randolph Blake. Hearing what the eyes see: auditory encoding of visual temporal sequences. *Psychological science*, 16(3):228–235, 2005. doi:10.1111/j.0956-7976.2005.00808.x.
- [164] Achille Pasqualotto and Michael J. Proulx. The role of visual experience for the neural basis of spatial cognition. *Neuroscience and biobehavioral reviews*, 36(4):1179–1187, 2012. doi:10.1016/j.neubiorev.2012.01.008.
- [165] Barbara G. Shinn-Cunningham. Object-based auditory and visual attention. *Trends in cognitive sciences*, 12(5):182–186, 2008. doi:10.1016/j.tics.2008.02.003.
- [166] Patrice Voss. Auditory Spatial Perception without Vision. *Frontiers in psychology*, 7:1960, 2016. doi:10.3389/fpsyg.2016.01960.
- [167] Scott A. Stone and Matthew S. Tata. Rendering visual events as sounds: Spatial attention capture by auditory augmented reality. *PLoS one*, 12(8):e0182635, 2017. doi:10.1371/journal.pone.0182635.
- [168] Raymond M. Klein. Covert Exogenous Cross-Modality Orienting between Audition and Vision. *Vision (Basel, Switzerland)*, 2(1), 2018. doi:10.3390/vision2010008.
- [169] Elena Aggias-Vella, Claudio Campus, Andrew Joseph Kolarik, and Monica Gori. The Role of Visual Experience in Auditory Space Perception around the Legs. *Scientific Reports*, 9(1), 2019. doi:10.1038/s41598-019-47410-2.

- [170] Hanne Stenzel, Jon Francombe, and Philip J. B. Jackson. Limits of Perceived Audio-Visual Spatial Coherence as Defined by Reaction Time Measurements. *Frontiers in neuroscience*, 13:451, 2019. doi:10.3389/fnins.2019.00451.
- [171] Xiaoxi Chen, Qi Chen, Dingguo Gao, and Zhenzhu Yue. Interaction between endogenous and exogenous orienting in crossmodal attention. *Scandinavian journal of psychology*, 53(4):303–308, 2012. doi:10.1111/j.1467-9450.2012.00957.x.
- [172] Voxelfarm. Voxel Farm - Voxel Engine and Procedural Generation in Unreal Engine and Unity, 3/13/2019. <https://www.voxelfarm.com/index.html>, Accessed on October 2020.
- [173] Mark Billingham, Adrian Clark, and Gun Lee. A Survey of Augmented Reality: Foundations and Trends® in Human-Computer Interaction, 8(2-3), 73-272. *Foundations and Trends® in Human-Computer Interaction*, 8(2-3):73–272, 2015. doi:10.1561/11000000049.
- [174] D. Brock, J. L. Stroup, and J. Ballas. Using an auditory display to manage attention in a dual task, multiscreen environment. In *International Conference on Auditory Display (ICAD)*. SMARTech, 2002. <https://smartech.gatech.edu/handle/1853/51394>.
- [175] James P. Corcoran. *Comparing the Effects of Mental Workload Between Visual and Auditory Secondary Tasks During Laparoscopy*. PhD thesis, Old Dominion University Libraries, 2019. doi:10.25777/KV21-7V83.
- [176] Hugo. Fastl and Eberhard. Zwicker. *Psychoacoustics: Facts and Models*, volume 22 of *Springer series in information sciences*. Springer Berlin Heidelberg and Imprint: Springer, Berlin, Heidelberg, 3rd ed. edition, 2007. doi:10.1007/978-3-540-68888-4.
- [177] Joseph Feliciano and Michael Stifelman. Robotic retroperitoneal partial nephrectomy: a four-arm approach. *JSLs : Journal of the Society of Laparoendoscopic Surgeons*, 16(2):208–211, 2012. doi:10.4293/108680812X13427982376149.
- [178] P. Ganesan, V. Rajini, B. S. Sathish, and Khamar Basha Shaik. HSV color space based segmentation of region of interest in satellite images. In *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT 2014)*, pages 101–105, [Piscataway, N.J.], 2014. IEEE. doi:10.1109/ICCICCT.2014.6992938.
- [179] Mari Riess Jones, Richard R. Fay, and Arthur N. Popper. *Music perception: Mari Riess Jones, Richard R. Fay, Arthur N. Popper*, volume 36 of *Springer handbook of auditory research, 0947-2657*. Springer, New York, 2010. doi:10.1007/978-1-4419-6114-3.
- [180] N. M. Gibbs and S. V. Gibbs. Misuse of 'trend' to describe 'almost significant' differences in anaesthesia research. *BJA: British Journal of Anaesthesia*, 115(3):337–339, 2015. doi:10.1093/bja/aev149.
- [181] R. Graham. Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. *Ergonomics*, 42(9):1233–1248, 1999. doi:10.1080/001401399185108.

- [182] Oleg Heizmann, Stephan Zidowitz, Holger Bourquain, Silke Potthast, Heinz-Otto Peitgen, Daniel Oertli, and Christoph Kettelhack. Assessment of intraoperative liver deformation during hepatic resection: prospective clinical study. *World journal of surgery*, 34(8):1887–1893, 2010. doi:10.1007/s00268-010-0561-x.
- [183] Javier Delgado del Hoyo. Multiple View Geometry | Real Time 3D Reconstruction from Monocular Video, 11/27/2015. <http://lfa.mobivap.uva.es/~fradelg/phd/tracking/multiview.html>, Accessed on October 2020.
- [184] Archie Hughes-Hallett, Erik K. Mayer, Hani J. Marcus, Thomas P. Cundy, Philip J. Pratt, Ara W. Darzi, and Justin A. Vale. Augmented reality partial nephrectomy: examining the current status and future perspectives. *Urology*, 83(2):266–273, 2014. doi:10.1016/j.urology.2013.08.049.
- [185] K. Alexander and M. Shansky. Influence of hue, value, and chroma on the perceived heaviness of colors. *Perception & Psychophysics*, 19:72–74, 1976.
- [186] Sanjeev Kaul, Rajesh Laungani, Richard Sarle, Hans Stricker, James Peabody, Ray Littleton, and Mani Menon. da Vinci-assisted robotic partial nephrectomy: technique and results at a mean of 15 months of follow-up. *European urology*, 51(1):186–91; discussion 191–2, 2007. doi:10.1016/j.eururo.2006.06.002.
- [187] Jon McCormack, Jonathan C. Roberts, Benjamin Bach, Carla Dal Sasso Freitas, Takayuki Itoh, Christophe Hurter, and Kim Marriott. Multisensory Immersive Analytics. In Kim Marriott, editor, *Immersive analytics*, volume 11190 of *LNCS sublibrary: SL 3 - Information systems and applications, incl. Internet/Web, and HCI*, pages 57–94. Springer, Cham, Switzerland, 2018. doi:10.1007/978-3-030-01388-2_{_}3.
- [188] Neil M. McLachlan. *Timbre, Pitch, and Music*, volume 1 of *Oxford Handbooks Online*. Oxford University Press, 2016. doi:10.1093/oxfordhb/9780199935345.013.44.
- [189] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: a class of displays on the reality-virtuality continuum. doi:10.1117/12.197321.
- [190] Stéphane Nicolau, Luc Soler, Didier Mutter, and Jacques Marescaux. Augmented reality in laparoscopic surgical oncology. *Surgical Oncology*, 20(3):189–201, 2011. doi:10.1016/j.suronc.2011.07.002.
- [191] George D. Ogden, Jerrold M. Levine, and Ellen J. Eisner. Measurement of Workload by Secondary Tasks: Human Factors: The Journal of the Human Factors and Ergonomics Society, 21(5), 529-548. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 21(5):529–548, 1979. doi:10.1177/001872087902100502.
- [192] Iqbal Singh. Robot-assisted laparoscopic partial nephrectomy: Current review of the technique and literature. *Journal of minimal access surgery*, 5(4):87–92, 2009. doi:10.4103/0972-9941.59305.

- [193] W. Dixon Ward. Chapter Eleven - Musical Perception. In Jerry V. Tobias, editor, *Foundations of modern auditory theory volume 1*, pages 405–447. Academic Press, S.l., 1970. doi:10.1016/B978-0-12-691901-1.50016-4.
- [194] Wikipedia. Computer-assisted surgery, 2020. https://en.wikipedia.org/wiki/Computer-assisted_surgery, Accessed on October 2020.
- [195] Leilei Xia, Xiaohua Zhang, Xianjin Wang, Tianyuan Xu, Liang Qin, Xiang Zhang, Shan Zhong, and Zhoujun Shen. Transperitoneal versus retroperitoneal robot-assisted partial nephrectomy: A systematic review and meta-analysis. *International journal of surgery (London, England)*, 30:109–115, 2016. doi:10.1016/j.ijssu.2016.04.023.
- [196] Tim Ziemer, Nuttawut Nuchprayoon, and Holger Schultheis. Psychoacoustic Sonification as User Interface for Human-Machine Interaction. <https://arxiv.org/pdf/1912.08609>.
- [197] Tim Ziemer, David Black, and Holger Schultheis. Psychoacoustic sonification design for navigation in surgical interventions. In *Proceedings of Meetings on Acoustics*, Proceedings of Meetings on Acoustics, page 050005. Acoustical Society of America, 2017. doi:10.1121/2.0000557.
- [198] Tim Ziemer and Holger Schultheis. Psychoacoustic auditory display for navigation: an auditory assistance system for spatial orientation tasks. *Journal on Multimodal User Interfaces*, 13(3):205–218, 2019. doi:10.1007/s12193-018-0282-2.

Appendix

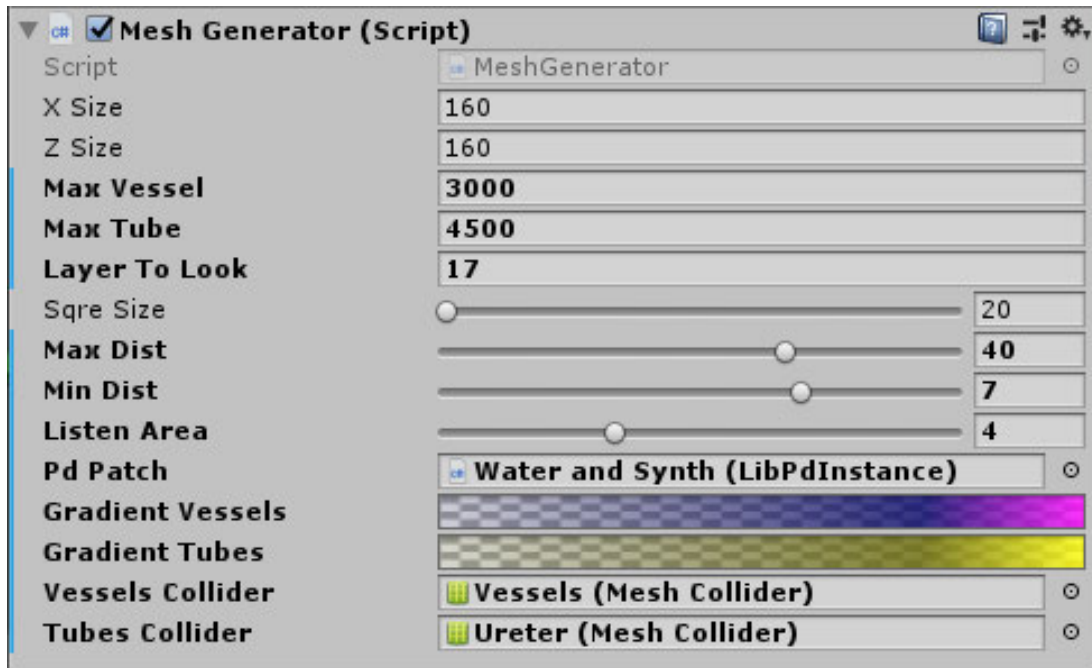


Figure A.1: Adjustable Variables for Pseudo-chromadepth Mesh.

The *X* and *Y Size* are the number of vertices in *X* and *Y*, respectively. *Max Vessel* and *Tube* represent the minimum number of vertices to label the structure as *high density*. *Sqre Size* is the dimensions of the mesh in mm^2 . *Max* and *Min Dist* are the maximum and minimum penetration depth, with which the colour values are normalised. *Listen Area* is the area of the listening zone in mm^2 . *Pd Patch* is the active auditory feedback design. The colour gradients are adjustable via *Gradient Vessels* and *Tubes*. The colliders for our structures are set via *Vessels* and *Tubes Collider*.

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

| Name | Task | Date |
|---|------|------|
| <p>Mental Demand</p> <p style="text-align: center;">Very Low Very High</p> <p>How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?</p> | | |
| <p>Physical Demand</p> <p style="text-align: center;">Very Low Very High</p> <p>How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?</p> | | |
| <p>Temporal Demand</p> <p style="text-align: center;">Very Low Very High</p> <p>How much time pressure did you feel due to the rate or pace at which the task elements occurred? Was the pace slow and leisurely or rapid and frantic?</p> | | |
| <p>Performance</p> <p style="text-align: center;">Perfect Failure</p> <p>How successful were you in accomplishing what you were asked to do? How satisfied were you with your performance in accomplishing these goals?</p> | | |
| <p>Effort</p> <p style="text-align: center;">Very Low Very High</p> <p>How hard did you have to work (mentally and physically) to accomplish your level of performance?</p> | | |
| <p>Frustration</p> <p style="text-align: center;">Very Low Very High</p> <p>How insecure, discouraged, irritated, stressed, and annoyed did you feel during the task?</p> | | |

Figure A.2: NASA-TLX Questionnaire.

Each scale represents one workload dimension with 21 ticks, counted from 0-20 or 1-21. The lower scores report lesser workload experienced by the user.

In the following 15 pairs, please select the scale title that represents the more important contributor to the workload of the task that you have performed.

- | | |
|--|--|
| <input type="checkbox"/> Effort | <input type="checkbox"/> Performance |
| <input type="checkbox"/> Performance | <input type="checkbox"/> Mental Demand |
| <input type="checkbox"/> Temporal Demand | <input type="checkbox"/> Performance |
| <input type="checkbox"/> Frustration | <input type="checkbox"/> Temporal Demand |
| <input type="checkbox"/> Temporal Demand | <input type="checkbox"/> Mental Demand |
| <input type="checkbox"/> Effort | <input type="checkbox"/> Effort |
| <input type="checkbox"/> Physical Demand | <input type="checkbox"/> Mental Demand |
| <input type="checkbox"/> Frustration | <input type="checkbox"/> Physical Demand |
| <input type="checkbox"/> Performance | <input type="checkbox"/> Effort |
| <input type="checkbox"/> Frustration | <input type="checkbox"/> Physical Demand |
| <input type="checkbox"/> Physical Demand | <input type="checkbox"/> Frustration |
| <input type="checkbox"/> Performance | <input type="checkbox"/> Mental Demand |
| <input type="checkbox"/> Temporal Demand | <input type="checkbox"/> Physical Demand |
| <input type="checkbox"/> Mental Demand | <input type="checkbox"/> Temporal Demand |
| <input type="checkbox"/> Frustration | |
| <input type="checkbox"/> Effort | |

Figure A.3: Weighted NASA-TLX Pairs.

The users select the most relevant workload dimension to the task based on their personal opinion. Only one dimension should be selected from each pair. This results in a total of 15 weights, distributed among six dimensions.

| | Distance | | | | Time | | | | Weighted TLX Rating | | | |
|--------------------|----------|--------|--------|--------|--------|--------|--------|--------|---------------------|--------|--------|--------|
| | NoA | VA | AA | VAA | NoA | VA | AA | VAA | NoA | VA | AA | VAA |
| Valid | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| Missing | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mean | 11.131 | 4.801 | 8.594 | 6.432 | 35.391 | 42.050 | 41.769 | 39.833 | 14.139 | 10.873 | 12.927 | 11.927 |
| Std. Error of Mean | 0.849 | 0.951 | 1.004 | 1.275 | 5.131 | 6.804 | 5.363 | 5.513 | 0.658 | 1.164 | 1.044 | 0.994 |
| Std. Deviation | 2.817 | 3.153 | 3.329 | 4.230 | 17.019 | 22.566 | 17.786 | 18.285 | 2.182 | 3.861 | 3.462 | 3.296 |
| Minimum | 6.027 | 0.992 | 3.963 | 1.884 | 11.169 | 17.117 | 23.290 | 13.437 | 10.800 | 6.000 | 6.267 | 7.200 |
| Maximum | 14.998 | 12.415 | 13.412 | 15.712 | 66.980 | 76.848 | 72.036 | 69.696 | 18.333 | 19.400 | 18.000 | 17.600 |

Table A.1: The Table of Descriptive Statistics.

The descriptive statistics can be observed in this table. For each assistance method, the statistics for the accuracy, time, and workload is provided.

| User ID | Assistance | Points | Assistance | Points | Assistance | Points | Assistance | Points |
|---------|------------|-------------------|------------|-------------------|------------|-------------------|------------|-------------------|
| P001 | AA | $C_2(4, 5, 6)$ | VAA | $C_4(10, 11, 12)$ | NoA | $C_1(1, 2, 3)$ | VA | $C_3(7, 8, 9)$ |
| P002 | VA | $C_3(8, 7, 9)$ | NoA | $C_2(5, 4, 6)$ | VAA | $C_1(2, 1, 3)$ | AA | $C_4(11, 10, 12)$ |
| P003 | NoA | $C_4(12, 11, 10)$ | AA | $C_3(9, 8, 7)$ | VAA | $C_1(3, 2, 1)$ | VA | $C_2(6, 5, 4)$ |
| P004 | VAA | $C_1(1, 2, 3)$ | NoA | $C_4(10, 11, 12)$ | VA | $C_2(4, 5, 6)$ | AA | $C_3(7, 8, 9)$ |
| P005 | VAA | $C_2(5, 4, 6)$ | VA | $C_1(2, 1, 3)$ | AA | $C_4(11, 10, 12)$ | NoA | $C_3(8, 7, 9)$ |
| P006 | NoA | $C_3(9, 8, 7)$ | VA | $C_4(12, 11, 10)$ | AA | $C_1(3, 2, 1)$ | VAA | $C_2(6, 5, 4)$ |
| P007 | VA | $C_4(10, 11, 12)$ | VAA | $C_3(7, 8, 9)$ | AA | $C_1(1, 2, 3)$ | NoA | $C_2(4, 5, 6)$ |
| P008 | AA | $C_1(2, 1, 3)$ | VA | $C_3(8, 7, 9)$ | VAA | $C_4(11, 10, 12)$ | NoA | $C_2(5, 4, 6)$ |
| P009 | AA | $C_4(12, 11, 10)$ | NoA | $C_3(9, 8, 7)$ | VA | $C_1(3, 2, 1)$ | VAA | $C_2(6, 5, 4)$ |
| P010 | VAA | $C_3(8, 7, 9)$ | AA | $C_2(5, 4, 6)$ | NoA | $C_4(11, 10, 12)$ | VA | $C_1(2, 1, 3)$ |
| P011 | NoA | $C_1(3, 2, 1)$ | VAA | $C_3(9, 8, 7)$ | VA | $C_4(12, 11, 10)$ | AA | $C_2(6, 5, 4)$ |
| P012 | VA | $C_2(4, 5, 6)$ | AA | $C_3(7, 8, 9)$ | NoA | $C_1(1, 2, 3)$ | VAA | $C_4(10, 11, 12)$ |

Table A.2: Evaluation Study's Assistance Methods and Points Progression.

Cluster one is located at the upper pole with points 1, 2, and 3. Cluster two is located in between the upper pole and the hilus, with points 4, 5, and 6. Cluster three is located in between the lower pole and the hilus, with points 7, 8, and 9. Cluster four is located at the lower pole with points 10, 11, and 12.

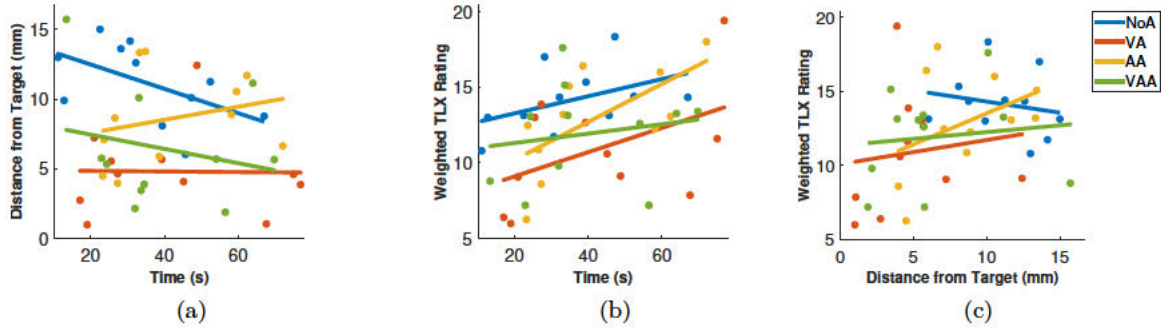


Figure A.4: Scatter Plots for the Evaluation Results.

In a) the points and their correlations based on the time-distant is plotted. b) depicts these points and correlations based on the time-weighted TLX rating. c) shows the correlations and points according to the distance-weighted TLX rating. The significance of this correlation is investigated in table A.3.

| | | Pearson | | Spearman | | Kendall | |
|-----------|---------------------|---------|--------|----------|--------|---------|--------|
| | | r | p | rho | p | tau B | p |
| Time- | Distance | | | | | | |
| | NoA | -0.528 | 0.095 | -0.536 | 0.094 | -0.382 | 0.121 |
| | VA | -0.015 | 0.966 | 0.045 | 0.903 | -0.018 | 1.000 |
| | AA | 0.250 | 0.458 | 0.336 | 0.313 | 0.273 | 0.283 |
| | VAA | -0.221 | 0.513 | -0.218 | 0.521 | -0.164 | 0.542 |
| Distance- | Weighted TLX Rating | | | | | | |
| | NoA | -0.194 | 0.568 | -0.237 | 0.482 | -0.167 | 0.481 |
| | VA | 0.133 | 0.697 | 0.382 | 0.248 | 0.236 | 0.359 |
| | AA | 0.403 | 0.219 | 0.336 | 0.313 | 0.273 | 0.283 |
| | VAA | 0.117 | 0.732 | 0.132 | 0.699 | 0.073 | 0.755 |
| Time- | Weighted TLX Rating | | | | | | |
| | NoA | 0.443 | 0.173 | 0.575 | 0.064 | 0.426 | 0.072 |
| | VA | 0.469 | 0.145 | 0.491 | 0.129 | 0.309 | 0.218 |
| | AA | 0.645 | *0.032 | 0.718 | *0.017 | 0.564 | *0.017 |
| | VAA | 0.173 | 0.612 | 0.369 | 0.264 | 0.257 | 0.274 |

* $p < .05$, ** $p < .01$, *** $p < .001$

Table A.3: Correlation Coefficients and significance.

This table investigates the correlation coefficients and significance for the linear models presented in fig. A.4. None of the correlations is significant, except the time-weighted TLX rating for the AA.