

# Media-enhanced cooking using a hands-free device

Evaluated through a HoloLens application

Adrian Föhl

A thesis submitted in partial fulfillment of the  
requirements for the degree of Master of Science

Stuttgart Media University  
Department of Computer Science and Media  
Faculty of Print and Media

First examiner: Prof. Dr. Simon Wiest  
Second examiner: Prof. Dr. Jürgen Scheible

May 29, 2018

## **Ehrenwörtliche Erklärung (Statutory declaration)**

---

Hiermit versichere ich, Adrian Föhl, ehrenwörtlich, dass ich die vorliegende Masterarbeit mit dem Titel:

**„Media-enhanced cooking using a hands-free device - Evaluated through a HoloLens application“**

selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden. Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§19 Abs. 2 Master-SPO der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Adrian Föhl

Aalen, 29. Mai 2018

## Abstract

Head Mounted Displays (HMD) are increasingly used in various industries. But apart from the industry environment, the potentials of HMDs in a private environment like at home has been relatively unexplored so far. What daily tasks can these help with, in the home kitchen for example?

The aim of this thesis is to obtain knowledge about the usefulness of such an HMD, the HoloLens, in combination with an application, while following a new recipe. Therefore a prototype application for the HoloLens got developed which guides a user through the cooking of a sushi burger by using multimedia content.

With a mixed method design, consisting of quantitative and qualitative methods, the HoloLens in combination with an application was evaluated by 14 participants.

Not only the weight of the device was a problem for users. The test also revealed that the display is darkening the view and participants tend to look below the glasses. An advantage is indeed to reach the next cooking step without the need of using hands and always having in sight what needs to be done next. Positive feedback was given as well for the application. Through voice control the user communicates to a character which will guide through the recipe by videos and text.

If in future the technical characteristics of HMD devices will improve, an application in this context will be of advantage in order to simplify learning a new recipe. This device, in combination with an application, could help early-middle stage cognitive impaired people and blind people to cook.

## Zusammenfassung

Datenhelme (Head Mounted Displays) kommen mehr und mehr zum Einsatz in der Industrie. Wie aber mögliche Einsatzgebiete in einer privaten Umgebung aussehen könnten, wurde bislang noch wenig erforscht. Welche Aufgaben könnten beispielsweise mittels eines Datenhelms im Küchenumfeld unterstützt werden?

Ziel dieser Masterthesis ist es, Einblicke zu bekommen wie eine Datenhelm, zusammen mit einer Anwendung, beim Kochen eines neuen Rezeptes unterstützen kann.

Dazu wurde eine Anwendung für die Microsoft HoloLens entwickelt, die einen Benutzer durch einzelne Schritte eines Sushi-Burgers leitet.

Mittels eines gemischten Forschungsdesigns, bestehend aus quantitativen und qualitativen Methoden, wurde die Datenbrille zusammen mit der Anwendung mit 14 Teilnehmern evaluiert.

Neben dem Beklagen des schweren Gewichtes der HoloLens, hat sich auch gezeigt, dass das Display die Sicht etwas verdunkelt und Benutzer deshalb dazu neigen unter die HoloLens zu schauen. Von Vorteil ist allerdings, dass man zum nächsten Kochschritt gelangt, ohne dabei seine Hände benutzen zu müssen und immer im Blick hat, welcher Schritt ausgeführt werden sollte. Positive Rückmeldungen gab es auch für die Anwendung. Über die Sprachsteuerung wird dabei wie mit Spielecharakter kommuniziert und durch das Rezept anhand von Videos und Text geführt.

Wenn sich in Zukunft die technischen Eigenschaften der Datenhelme verbessern, könnte eine Anwendung in diesem Bereich von Vorteil sein, um das Erlernen von neuen Rezepten zu vereinfachen. Der Datenhelm, zusammen mit einer Anwendung, könnte kognitiv eingeschränkten Menschen im mittleren Stadium sowie blinden Menschen beim Kochen unterstützen.

# Contents

List of Figures	viii
List of Tables	ix
Acronyms	x
<b>1 Introduction</b>	<b>1</b>
1.1 Context of the Study . . . . .	1
1.2 Statement of the Problem . . . . .	2
1.3 Aim and Scope . . . . .	3
1.4 Significance of the Study . . . . .	4
1.5 Overview of the Study . . . . .	4
<b>2 Basics</b>	<b>5</b>
2.1 Augmented Reality . . . . .	5
2.1.1 Definition Of AR . . . . .	5
2.1.2 Differentiating With Other Realities . . . . .	6
2.1.3 Possible Areas for the Use Of AR . . . . .	7
2.2 Head-Mounted Display . . . . .	8
2.3 Microsoft HoloLens . . . . .	8
2.3.1 HoloLens Specification . . . . .	9
2.3.2 Interaction With the Device . . . . .	11
2.3.3 Limitations Of the Device . . . . .	11
2.3.4 Virtual Objects . . . . .	12
2.4 SmartKitchen Project . . . . .	12
2.5 Conclusion . . . . .	13
<b>3 Prior Work</b>	<b>14</b>
3.1 Proof Of Concept Applications . . . . .	14
3.2 Guided Assistance Through a HoloLens . . . . .	15
3.3 Cooking Instructions in the Kitchen . . . . .	15
3.4 SmartKitchen Research Project . . . . .	16
3.5 Conclusion . . . . .	17

---

<b>4</b>	<b>Application</b>	<b>18</b>
4.1	Requirements . . . . .	18
4.1.1	Interaction Within the Application . . . . .	18
4.1.2	Character Characteristics . . . . .	19
4.1.3	Cooking Guidance . . . . .	20
4.1.4	Recipe Recommendation . . . . .	20
4.2	Design and Technology Choices . . . . .	21
4.2.1	Character . . . . .	21
4.2.2	Recipe . . . . .	22
4.2.3	Image Recognition . . . . .	23
4.2.4	Language . . . . .	24
4.2.5	Tools . . . . .	24
4.3	System Overview . . . . .	25
4.4	Implementation . . . . .	26
4.4.1	Backend Services . . . . .	26
4.4.2	Frontend . . . . .	30
4.4.3	Challenges . . . . .	33
4.5	Conclusion . . . . .	34
<b>5</b>	<b>Method</b>	<b>35</b>
5.1	Research Design . . . . .	35
5.2	Instruments . . . . .	36
5.3	Participants . . . . .	37
5.4	Procedure . . . . .	38
5.5	Data Analysis Plan . . . . .	41
5.6	Legitimation . . . . .	42
5.7	Conclusion . . . . .	42
<b>6</b>	<b>Results</b>	<b>43</b>
6.1	Participation Profiles . . . . .	43
6.2	HoloLens . . . . .	44
6.2.1	Hands-free . . . . .	44
6.2.2	Optics . . . . .	45
6.2.3	Comfort . . . . .	45
6.2.4	Interaction . . . . .	46
6.2.5	Overall Remarks . . . . .	46
6.3	Application . . . . .	47
6.3.1	Character . . . . .	47
6.3.2	Video and Text . . . . .	49
6.3.3	Scanning Feature . . . . .	51
6.3.4	Additional Features . . . . .	51
6.3.5	Results About the Overall Application . . . . .	52
6.4	Conclusion . . . . .	54

---

<b>7</b>	<b>Discussion</b>	<b>55</b>
7.1	Evaluation Of the HoloLens . . . . .	55
7.2	Evaluation Of the Application . . . . .	56
<b>8</b>	<b>Conclusion and Future Work</b>	<b>59</b>
8.1	Conclusions . . . . .	59
8.2	Future Work . . . . .	60
<b>A</b>	<b>Cooking Flow</b>	<b>62</b>
A.1	Welcome Flow . . . . .	62
A.2	Cooking Inspiration . . . . .	62
A.3	Cooking Guide . . . . .	63
A.4	Goodbye Message . . . . .	64
<b>B</b>	<b>Questions</b>	<b>65</b>
B.1	Semi-structured Interview . . . . .	65
B.1.1	During Salmon Cooking . . . . .	65
B.1.2	After the Sushi Burger Is Done . . . . .	65
B.2	Questionnaire . . . . .	65
B.2.1	Background Information . . . . .	65
B.2.2	Questions Regarding Application . . . . .	66
B.2.3	Questions Regarding HoloLens . . . . .	66
	<b>Bibliography</b>	<b>67</b>

# List of Figures

2.1	Virtuality continuum concept (Adapted from: Milgram and Kishino (1994)) . . . .	6
2.2	Microsoft mixed reality spectrum (Adapted from: Bray and Zeller (2018)) . . . .	6
2.3	Augmented reality investment by industry (PwC (2017)) . . . . .	7
2.4	Microsoft HoloLens device . . . . .	9
2.5	Human Centered Design (International Organization for Standardization (Mar 2010))	13
4.1	Microsoft Clippy character . . . . .	19
4.2	The final sushi burger . . . . .	23
4.3	System overview of the application . . . . .	26
4.4	Microsoft Bot Framework sources in Microsoft Azure . . . . .	27
4.5	Microsoft LUIS with an example intent . . . . .	28
4.6	Training set for avocados in Microsoft Custom Vision . . . . .	30
4.7	Character of the application . . . . .	30
4.8	Recipe selection in the application . . . . .	32
4.9	A cooking step in the application . . . . .	33
5.1	Kitchen table setup . . . . .	38
5.2	Video recording view of a participant . . . . .	39
6.1	Reasons for not cooking more often . . . . .	43
6.2	Where participants learned to cook . . . . .	44
6.3	Character or voice preference . . . . .	47
6.4	Text and character overlapping . . . . .	50
6.5	AttrakDiff - Portfolio diagramm . . . . .	52
6.6	AttrakDiff - Diagram of average values . . . . .	52
6.7	AttrakDiff - Description of word-pairs . . . . .	53



# List of Tables

- 4.1 Comparison of different object recognition services . . . . . 23
  
- 5.1 Used instruments overview . . . . . 36
- 5.2 Sushi burger recipe . . . . . 39
  
- 6.1 Optic problems overview . . . . . 45
- 6.2 Comfort issues overview . . . . . 46
- 6.3 Triggered emotions findings . . . . . 48
- 6.4 Video issues findings . . . . . 50
- 6.5 Text issues findings . . . . . 51

# Acronyms

AR augmented reality.

FOV field of view.

HCD human-centered design.

HMD head-mounted display.

IoT Internet of Things.

LUIS Language Understanding Intelligent Service.

MR mixed reality.

POC proof of concept.

RQ research question.

VR virtual reality.

# Introduction

## 1.1 Context of the Study

Since 2017 developers have created a variety of applications for iOS and Android platforms where reality is enhanced with additional information, from displaying simple text to complex 3D models. This enhancement is better known as augmented reality (AR). Augmented reality isn't something new, it has already been around for a long period of time but regardless it's still considered to be in early development (Azuma et al. (2001)). Only since Apple introduced the ARKit<sup>1</sup> library in June 2017, and Google launched ARCore<sup>2</sup> in August 2017, has AR been hyped through the media (Park (2017); Bardi (2017); Buche (2017); Gorey (2017)). It seems that technology is at a point where it can now use the full potential of AR.

Whereas most AR applications are developed for mobile devices, a small but growing number of applications are targeted for wearables devices like head-mounted display (HMD). An HMD projects visuals on a display screen attached to your head or as part of a helmet. The main advantage of such a device is that it leaves hands free to do other things. Especially in the industry environment more and more AR use cases are explored and developed to increase workers' productivity (Annunziata and Abraham (2017)). Workers don't have to switch context between their actual tasks and retrieving the information they need. Information is displayed right directly in front of them.

Unlike an HMD on the other hand, smart glasses can only enhance reality on a basic level. Smart glasses can't display information in a way that feels that the information is integrated into the real world. The best example for smart glasses is Google Glass, introduced to the public in April 2012. It was equipped with a touchpad, allowing interaction on the side of the Google Glass, a camera, taking up to 5 Megapixels, and a small display screen (Google (n.d.)). Among the applications were taking pictures and videos, navigation, displaying search results and voice translation. However, it created a lot of concern regarding privacy rights to the point that some establishments like cinemas started to ban Google Glass (Sherwin (2014)). Google decided in January 2015 to stop the prototype production (Google Glass (2015)).

In 2017 a second generation was announced as an enterprise edition. Between 2015 and 2017 Google still worked on Google Glass but it was only available to a mere fifty companies through a special program. Thus, the Google Glass has been thoroughly tested with companies like DHL,

---

<sup>1</sup><https://developer.apple.com/videos/play/wwdc2017/602/>

<sup>2</sup><https://developers.google.com/ar/discover>

General Electrics and Boing (Palladino (2017)). The positive feedback led to the opening up of the program in order to gain more partners.

But apart from the industry environment, the potentials of AR and smart glasses or HMD in a private environment like at home has been relatively unexplored so far. Is there a need for a hands-free device and application for the home? What daily tasks can these help with, in the home kitchen for example?

At present there is a trend for everything to be smart. Smart homes, smart speakers and smart toothbrushes for example. There is also the smart kitchen, which includes smart cooking appliances. Based on a study in 2015 from GFK (Heyder (2016)), sales of kitchen machines increased in Western Europe by 3.2% from 2014. Probably, one of the most hyped kitchen devices in recent years in Germany is the Thermomix TM 5 produced by Vorwerk (Welzhofer (n.d.)). It is an all-in-one device for the kitchen. It can blend and mix ingredients, as well as weigh, chop and heat (*Thermomix – Unser Multitalent* (n.d.)). It also offers guided cooking, showing each step of a chosen recipe on a screen. At a price of 1,199€, the Thermomix TM 5 is not cheap. Regardless of the retail price, in 2015 the company had a turnover of 49% compared to 2014. In 2016 sales increased by 11% and in 2017 it decreased by 12.9% according to Vorwerk sales report (*Publikationen* (n.d.)). The drop in sales doesn't mean that such a device is not in demand any more. Competitors released similar products and a whole range of devices are now in the market.

Another innovation is coming from the startup HelloFresh. If someone is lacking in inspiration regarding what to cook, they offer a solution. They send all ingredients for a meal along with a recipe (HelloFresh (n.d.)). In 2016 alone the company made a turnover of 95% increase in revenue compared to 2015, or 597 million euros (Rocket Internet (2017)). In 2017 it made more than 900 million euros (HelloFresh (2018)). Based on those numbers, it seems that people are looking for help in the kitchen.

From a smart fridge<sup>3</sup> to a smart spoon<sup>4</sup>, everything can be found nowadays. But would AR and a smartglass/HMD make sense in the cooking context? The present thesis will explore how an HMD, in particular the HoloLens, can support in the home kitchen.

## 1.2 Statement of the Problem

The earliest recipes were found on three clay tablets dating back to 1700 BC from the ancient Mesopotamians (Bottéro (1985)). Since then cooking with a recipe has barely changed. Cookbook recipes simply became enhanced with meaningful pictures.

With the invention of the television, cooking shows started to be broadcasted. In 1946, Philip Harben was the first person to have a cooking show on television (Rosemary (2018)). With the introduction of the Internet, websites presenting all kind of different recipes from all over the world have appeared. There is now a growing number of video clips about cooking, hosted on different platforms, from both professional and amateur chefs.

In addition, advances in technology has made an impact in the process of cooking. These days, table computers and smartphones are being used in the kitchen for reading recipes. But some people still go to cooking websites in the Internet and print out recipes to take to the kitchen.

---

<sup>3</sup>A fridge that keeps track what is inside and has internet connection

<sup>4</sup>A spoon which stabilize handles

A study was conducted in September 2016 by Forsa Germany asking 1,200 people (TK (2016)) for the reasons they don't cook. Besides the main reason that 49% of the people don't have time, it also revealed that 17% do not know how to cook. According to statistics (IfD Allensbach (2017)), the number of people between 2013 and 2017 who never cook in Germany decreased by 3.8%. But it is still high at 12.4 million people. That more people are cooking may be due to several reasons such as following a diet, wanting to eat healthier or having allergies.

Cooking with a recipe has always been problematic though. Switching between reading the recipe, be it on a book or some device, and the actual cooking is time-consuming and error-prone, particularly in keeping track of where the next step is on the page. In addition, turning to a next page or step can get unhygienic because hands are needed to interact with the physical object, not to mention get the book or device dirty.

To date, however, there is no research that examines if a smart glass or an HMD can assist in the cooking process. If using such a device indeed makes a difference while following a recipe, it is equally important to understand what device restrictions exist and how an application could assist. In summary, if this research could add all the pieces of information together it would be quite intriguing to know the degree to which the use of such a device in the kitchen can assist.

### 1.3 Aim and Scope

The aim of this thesis is to obtain knowledge about the usefulness of a HoloLens, in combination with an application, while following a new recipe. Hence quantitative and qualitative methods are used to identify the factors which influence the usage of such a device.

The following questions are set to answer this goal:

RQ1: Which restrictions currently exist in using a HoloLens device in the kitchen context?

RQ2: In what way can a hands-free device, coupled with media support, assist while following new recipes?

To find out the answer to these questions, a prototype application was developed. Limitations to the research were given from the start. For this present study, the chosen recipe spanned around one hour. This duration was long enough that the person cooking was immersed in the activity.

This study is not about examining the person's cooking skills. Yet, participating testers should have average cooking skills. The research is focusing on a person's perception with a HoloLens while cooking. It will not focus on comparisons between cooking with a HoloLens and without, nor comparisons among cooking different kinds of recipes. For this reason, a recipe which can be done with an average knowledge of cooking was chosen. All cooking were conducted in a kitchen setup at university. Therefore, all participating testers were students.

In this study, the developed application was limited to basic functionality in terms of interaction and animation. The scope was not to implement a fully functional character which is capable of complex conversations. The application should guide through the recipe to achieve the end result. Due to language limitations the interaction in the application was only possible in English. Hence, all testers needed a basic level of English language skills.

Further, the aim is not to propose improvements to the HoloLens. The research will focus on the features that can potentially make people use such a device for cooking, and noticeable restrictions or limitations of the device.

## 1.4 Significance of the Study

The main intended outcome of the study is to examine the usefulness of an HMD like the HoloLens while following recipes. Specifically, focusing on cooking amateurs at the age of twenty-something cooking something new. Overall, the study will investigate if the usage of the HMD in the kitchen is a feasible use case at all. A second intended outcome, on a practical level, is to identify if there is a need for the necessary information to be displayed on a hands-free device during the cooking process and how such information could be presented in an application. Evidence from observation and interview questions will give a deeper insight. The last intended outcome of this study is to examine if the given step-by-step cooking instructions from the application contributes to an easier way of cooking.

## 1.5 Overview of the Study

This thesis consists of eight chapters divided into three main parts.

Chapter 2 and 3 make up Part I, providing background information and prior work. Chapter 2 will provide basic background information mainly about AR and the HoloLens. Chapter 3 will discuss similar works and studies, showing what knowledge exists in this area so far.

Part II consists of Chapter 4 and Chapter 5. Chapter 4 describes the application developed for the HoloLens, from requirements up to implementation, whereas Chapter 5 gives insights on how the test for this application was conducted and which research design was used.

In Part III (Chapters 6, 7 and 8) the results are provided, with a discussion and then conclusions about the cooking sessions carried out. Chapter 6 reflects on the results of the cooking. In Chapter 7, discussions of the key findings are developed. And the last chapter, Chapter 8, contains the conclusion and ideas for further research in this area.

# Basics

Now that the problem has been stated, basic knowledge is needed to process this research. Therefore, this chapter will provide some basic background information. The focus will be on AR, what it is and how it is used, followed by detailed information on a device which enables AR, the HoloLens.

## 2.1 Augmented Reality

In 2017 the term AR began appearing everywhere. One reason for the hype was the annual Worldwide Developers Conference (WWDC) keynote<sup>1</sup> by Apple introducing ARKit which took place on June 5 in San Francisco. But what is behind the term AR? The next section will explain AR in detail.

### 2.1.1 Definition Of AR

The term AR garnered people's attention when a smartphone game called Pokémon Go got released by the company Niantic on July 6, 2016. In this game, Pokémons are overlaid on the view of the real world as if Pokémons are integrated in the real surrounding. But there is a controversy on whether this game is truly using AR technology (Schart and Tschanz (2018)). It is argued that the game only shows the basic idea behind AR and is rather more of a location-based game. Nevertheless, it is estimated that this game has been downloaded around more than 750 million times as of June 2017 (Niantic (2017)).

The most commonly used definition of AR is from Azuma (1997): "AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it." Azuma defines three characteristics which an AR system needs to have:

1. Combines real and virtual
2. Interactive in real time
3. Registered in three dimensions

---

<sup>1</sup><https://www.apple.com/lae/apple-events/june-2017/>

But according to Papagiannis (2017), the definition of Azuma is missing one keyword: context. This transforms AR to the next level. Instead of giving an overall experience for everyone, the experience will be specific for the individual based on location, interests and needs. AR will have even deeper impact once several technologies are combined with AR such as Internet of Things (IoT), machine learning (ML) and artificial intelligence (AI).

### 2.1.2 Differentiating With Other Realities

Frequently used in connection with AR are terms like virtual reality (VR) and mixed reality (MR). But a common understanding for some terms is still deceptive. For a definition of terms, most people and companies refer to the virtuality continuum concept by Milgram and Kishino (1994).

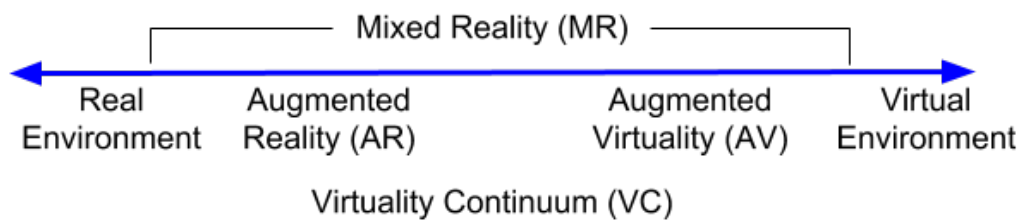


Figure 2.1: Virtuality continuum concept (Adapted from: Milgram and Kishino (1994))

In this figure 2.1 two different environments exist. At one end the real environment, and at the other end, the virtual environment, wherein the real environment is the physical world and the virtual environment is a world which is artificial and immersive to the user. VR uses an HMD which can't be seen through and is not part of MR, according to Milgram and Kishino (1994). MR is situated in between and combines the AR and augmented virtuality (AV) world. In AV, objects of the real world are placed into the virtual world.

Microsoft is responsible for much of the confusion. For simplicity's sake, Microsoft created their own MR spectrum (see figure 2.2 below). According to Fink (2018), the Microsoft Milgram scale lets people assume that VR is an extension of AR. But they are doing completely different things. To the author, VR is a new reality, whereas AR is an enhancement of reality. Based on the Microsoft figure 2.2, AR and VR belong to MR. For this thesis the term MR will not be used.

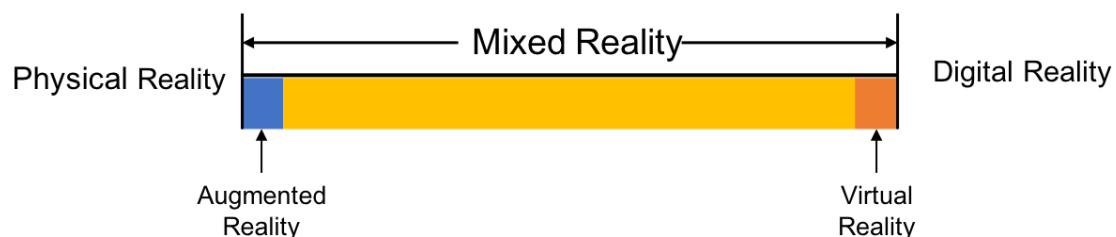


Figure 2.2: Microsoft mixed reality spectrum (Adapted from: Bray and Zeller (2018))



### 2.1.3 Possible Areas for the Use Of AR

AR has several advantages in all kinds of fields. Figure 2.3 shows a study from Pricewaterhouse-Coopers (PwC) which indicates how the various industries is currently investing in AR and how they will invest in the next three years.

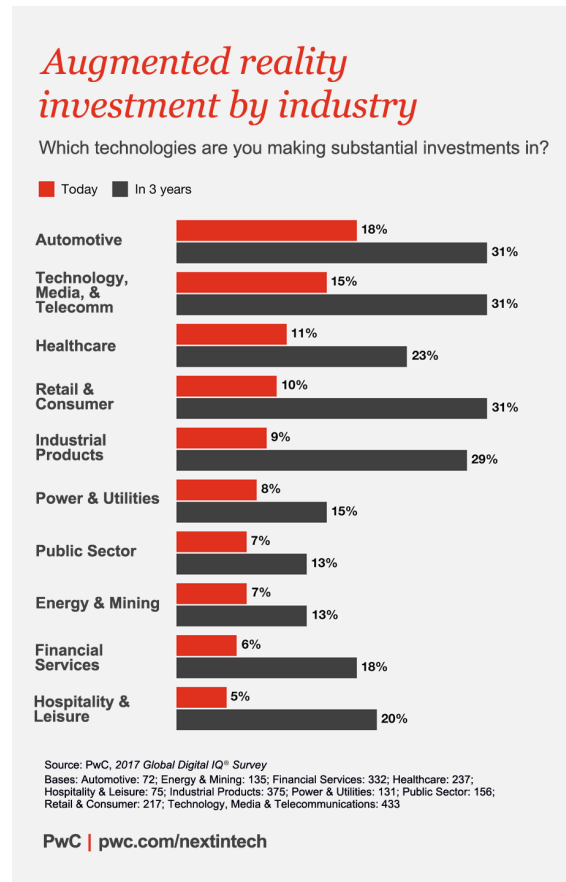


Figure 2.3: Augmented reality investment by industry (PwC (2017))

Notice that the PwC study states that AR will influence every industry, and not insignificantly. The ways it will influence is discussed by Porter and Heppelmann (2018). According to them, AR is adding value in three different ways: visualization, guidance and interaction.

#### Visualization

AR makes you see things which you can't see with the normal eye, adding some kind of x-ray view of products by overlaying a 3D model of an object to the real object.

#### Guidance

The user get guided by an AR application step-by-step, for example for constructing something. Or complex 2D objects are represented as interactive 3D objects, making it easier to understand things.

## Interaction

Instead of using physical control elements, AR enables the user to replace them through virtual elements, which are projected on the target.

## 2.2 Head-Mounted Display

An HMD projects visuals on a display attached to a person's head or part of a helmet. HMD can be used for AR and VR. They either have a display in front of one (monocular HMD) or two eyes (binocular HMD) (Lippert (1990)). HMD isn't something new. Often "The Sword of Damocles" created by Sutherland (1968), is cited as the first AR/VR HMD device. The device was limited in terms of functionality and realism and it wasn't fully immersive. The user merely saw geometric shapes which moved in sync with the movement of the head. The device was also so heavy that it was attached to the ceiling.

Technology has improved since then and HMD devices exist in several forms and for different industries. Some HMD just project images on the display while other devices can track position, angle and orientation of the head, as well as eye and hand movement. There are basically two types of HMD which exist (Rolland and Fuchs (2000)):

**See-through display** A user can look through a transparent display and see the real world while images are projected on the display.

**Video-through display** A video feed from the camera is showing content on a display. The user is not able to see through the display.

This thesis focuses on see-through HMD devices since the real world is still visible. The most notable devices as of now are: Microsoft HoloLens, Meta 2 and Magic Leap One. The Magic Leap One is only supposed to ship by end of 2018 and not a lot of information is available regarding the device, but it gets a lot of hype because the company was able to attract a lot of investors (Magic Leap (2018)). Investment has reached \$2.3 billion as of March 2018 (Spangler (2018)).

Meta 2 is already available with a price of 1,683€ (Meta Company (2018)). It has a 90-degree field of view (FOV), 2.5K resolution, 720p front-facing RGB camera and weighs 500g. It also supports the use of the grab gesture, which enables the user to naturally interact with objects. The downside of this device is that the user needs to be attached to a PC with a cable. In a kitchen context, wandering around with a cable could cause some issues. For this reason a HoloLens device seems to be the best choice.

## 2.3 Microsoft HoloLens

During a Windows 10 consumer preview in January 2015, Microsoft unveiled their first HMD device (Savov (2015)), see figure 2.4. It was still a prototype at the time and only got released to the public a year later in March 30, 2016.

The device is sold in two editions<sup>2</sup>. The developer version costs 3,299.00€ and is only for developers. The other version is the commercial edition, which is for enterprise developers and

<sup>2</sup><https://www.microsoft.com/en-us/hololens/buy>

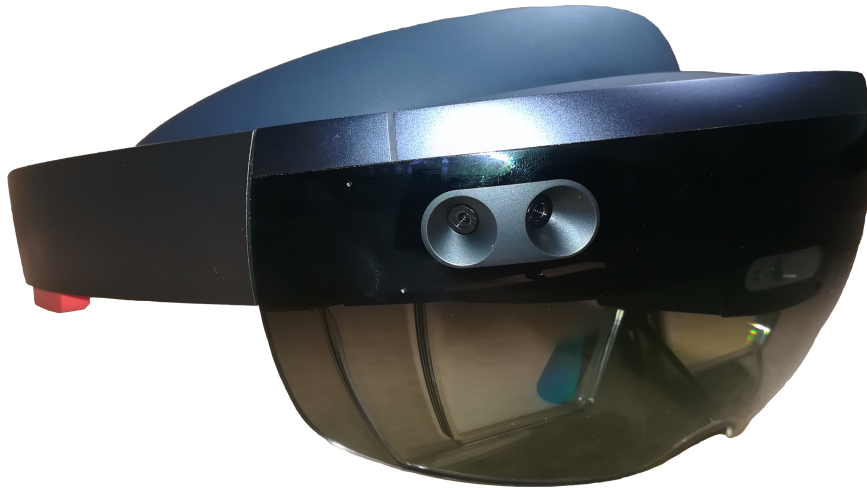


Figure 2.4: Microsoft HoloLens device

costs 5,489.00€. As of today Microsoft has not released any further versions of this device. The speculation is that Microsoft will skip version 2 and go directly to version 3, which will be released sometime in 2019 according to rumours on the Internet (Sams (2017)).

Meanwhile the price tag does not make it affordable for a normal consumer. And the consumers who do buy the device are bound to be disappointed because only a handful of things can be done with it currently. There is still a lack of applications which would justify the need for it for a normal consumer. Kipmann (2017) stated that around 150 applications exist for the device as of March 2017.

In a video presentation from European Patent Office (2018) the narrator says: "Microsoft says that about 50,000 copies have been sold so far, but many believe in the potential of these smart glasses." However, there is no confirmation from Microsoft so far that this number is correct. Several companies like NASA's Jet Propulsion Laboratory, Ford, Thyssenkrupp and many others experimented with the device, but how it is used in daily work is hard to tell (*Transform your business* (n.d.)).

The operating system Windows 10 runs on the HoloLens, which makes it a standalone device. Through the operating system a user can access the application store to install applications. It contains a powerful holographic processing unit (HPU) which is responsible for analysing information like mapping the environment, or recognizing gestures from all available sensors.

### 2.3.1 HoloLens Specification

Zeller and Bray (2018) gives a detailed specification of the HoloLens device, these specifications are listed below.

#### Optics

- See-through holographic lenses (waveguides)
- 2 HD 16:9 light engines
- Automatic pupillary distance calibration
- Holographic Resolution: 2.3M total light points

- Holographic Density: >2.5k radiants (light points per radian)

### Sensors

- 1 Inertial Measurement Unit (IMU)
- 4 environment understanding cameras
- 1 depth camera
- 12MP photo / HD video camera
- Mixed reality capture
- 4 microphones
- 1 ambient light sensor

### Human Understanding

- Spatial sound
- Gaze tracking
- Gesture input
- Voice support

### Input / Output / Connectivity

- Built-in speakers
- Audio 3.5mm jack
- Volume up/down
- Brightness up/down
- Power button
- Battery status LEDs
- Wi-Fi 802.11ac
- Micro USB 2.0
- Bluetooth 4.1 LE

### Power

- Battery Life
  - 2-3 hours of active use
  - Up to 2 weeks of standby time
  - Fully functional when charging
- Passively cooled (no fans)

### Processors

- Intel 32 bit architecture with TPM 2.0 support
- Custom-built Microsoft Holographic Processing Unit (HPU 1.0)

### Weight

- 579g

### Memory

- 64GB Flash
- 2GB RAM

## 2.3.2 Interaction With the Device

Because the device is new, a new way of interaction is needed. To interact with the HoloLens and its content, three possibilities exist: gaze, gestures and voice.

### Gaze

With the HoloLens on the user's head, the device always knows where and what the user is looking at in space, based on the position and orientation (Turner et al. (2018)). To give feedback to the user, an indicator is shown, normally a small dot but can also be anything else. This is similar to a mouse cursor which tells the system what the user is aiming at.

### Gesture

To reach out within the space for an object, hand gestures are needed. As of now two different core gestures are recognized: Air tap and Bloom (Rwinj et al. (2018a)).

**Air tap** Once an object is targeted through gaze, an interaction can be triggered by pressing two fingers together. It is recommended to raise the index finger up and then move it down to the thumb to press it together.

**Bloom** Whenever the user wants to go back to the main menu, a bloom gesture needs to be performed. Therefore the user needs to show a palm, facing up, while holding all finger tips together with the hand open.

### Voice

Instead of using a gesture to complete an interaction, voice commands can be used. Therefore voice commands need to be implemented within the application. For a better understanding, Microsoft gives the opportunity for the developer to add "voice dwell tips" (Rwinj et al. (2018b)). When a button is in focus, a tip will appear below the object, saying which voice command can be used to trigger it.

## 2.3.3 Limitations Of the Device

The FOV is limited to 30° horizontal and 17.5° vertical (16:9) (Kreylos (2015)). Humans typically have horizontal 180° FOV when facing forward and about 150° vertical (Mazuryk and Gervautz (1999)). When looking through a HoloLens to view objects that are big and the distance to the object is small, the object will be cut off. This effect is called clipping.

Today, HMD devices still have the big drawback in that they can operate only for a limited time. For the HoloLens it's two to three hours of active use. Using the device for a whole work shift would not be possible as the HoloLens must be recharged several times. Another option would be to plug in the device while in use but that defeats the purpose of an HMD device. Another possibility would be the use of a portable powerbank<sup>3</sup> connected to the HoloLens while in use. The downside is that there is more things for the user to carry. The problem is that the batteries are built in and can't be changed.

### 2.3.4 Virtual Objects

Beau et al. (2018) calls "objects made of light and sound" which are seen through the HoloLens, holograms. However, the definition of a hologram clearly states that a hologram can be seen without the need of any additional device. Oxford Dictionary (2018) defines it as "a three-dimensional image formed by the interference of light beams from a laser or other coherent light source." It is for this reason that this thesis refers to holograms as virtual objects.

Virtual objects can be placed at a fixed position in the application world while using the HoloLens, making it possible to walk around the object as if they were actually placed in the real world. On the other hand, virtual objects can follow your movement and can be positioned relative to you. The position of the virtual object in the device is best in a distance between 1.25m and 5m according to Beau et al. (2018). Otherwise an effect like clipping (see 2.3.3 Limitations) or fading out of the content will be noticeable in the HoloLens, which will reduce the immersive experience.

Due to the fact that virtual objects are added through light to a scene in an application, it's not possible to add black objects as light cannot be removed. As a result, black will appear always transparent in a HoloLens application.

## 2.4 SmartKitchen Project

Scheible et al. (2016) defines the aim of smartKitchen project as "developing solutions for integrating digital and personalized media into the future cooking environment. These solutions aim at making the cooking process more intuitive, social and fascinating."

SmartKitchen is a research project conducted at Stuttgart Media University funded by the German Federal Ministry of Education & Research (*smartKitchen project* (n.d.)). This project started in February 2016 and will last until February 2019. This project is headed by Prof. Dr. Jürgen Scheible and his associate leader Prof. Dr. Arnd Engeln from the faculty of Electronic Media. Involved are people from different departments which cover user research, implementation and data privacy. Furthermore, synergies are used together with project partners E.G.O. Elektrogerätebau Oberderdingen GmbH and Institute for Visualization and Interactive Systems (VIS) at University of Stuttgart.

The SmartKitchen project is following the human-centered design (HCD) for interactive systems also known as ISO 9241-210:2010. In order to guarantee that expectations and needs of the user are taken into consideration. International Organization for Standardization (Mar 2010) introduces HCD as following: "Human-centred design is an approach to interactive systems development that aims to make systems usable and useful by focusing on the users, their needs

---

<sup>3</sup>A portable power supply device

and requirements, and by applying human factors/ergonomics, and usability knowledge and techniques. This approach enhances effectiveness and efficiency, improves human well-being, user satisfaction, accessibility and sustainability; and counteracts possible adverse effects of use on human health, safety and performance." The process is seen in figure 2.5.

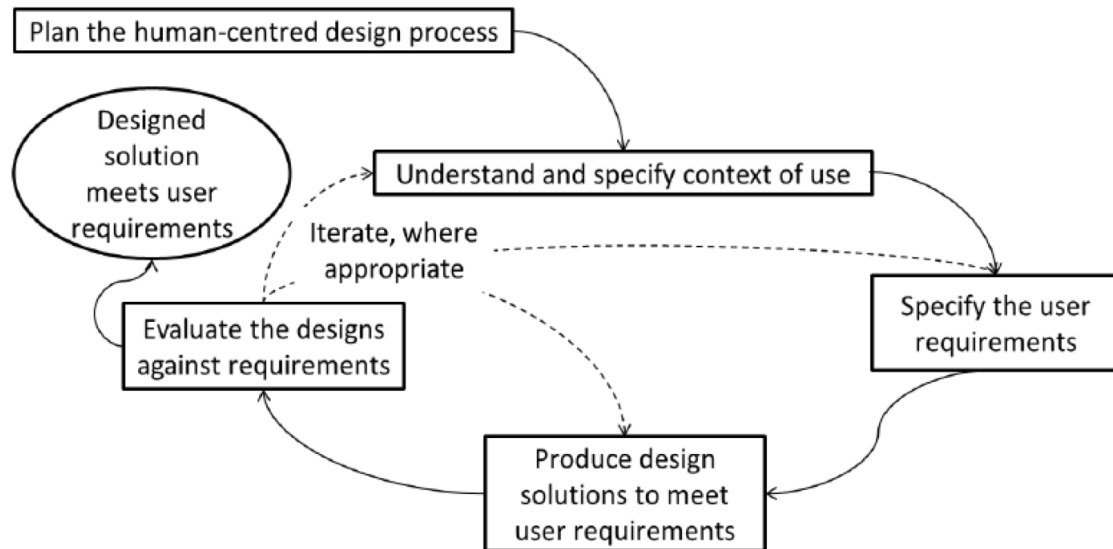


Figure 2.5: Human Centered Design (International Organization for Standardization (Mar 2010))

For the smartKitchen project a detailed heterogeneous user study was conducted to observe and analyse people's cooking behaviour, as well as to identify positive or negative user experience of tasks while cooking. In figure 2.5 this is the step called "Understand and specify context of use."

From the collected information, opportunity areas were extracted which specified user requirements. In the next step, ideation phase, prototypes were developed and tested. When needed, the prototypes were iterated, and user requirements were adjusted and tested again. Within the project, so called "in-situ projections" were tested, meaning that cooking-related content was projected on any surface in the kitchen where needed.

## 2.5 Conclusion

As mentioned earlier in this chapter, AR enables enhancement of reality with information on an as needed basis. This additional information when coupled with an HMD device like the HoloLens, allows for seamless display over the real world. This leaves both hands free to do other tasks. In a kitchen context this is enormous help. The chapter which follows will look into prior and similar research projects to understand how this combination of AR and HMD can help in the kitchen.

## Prior Work

The previous chapter presented details about terms, concepts and hardware. The goal for this chapter is to provide similar works and research in the field of cooking and HMD devices.

While a lot of research has been conducted on HMDs, AR or cooking, no research has been carried out to discover the usage of an HMD device in the kitchen context yet. However, three proof of concepts (POCs)<sup>1</sup> were found where a HoloLens application was developed for the use in the kitchen. These projects will be explained shortly. In addition previous research papers regarding the general use of HMD devices will be discussed as well as research regarding cooking.

### 3.1 Proof Of Concept Applications

The specific context "HMD device in the kitchen" seems not to have been the focus of researchers so far. Nevertheless, two university capstone projects<sup>2</sup> and one Hackathon project that deal with this topic exist.

#### **HoloCook**

HoloCook, a project from Nelson et al. (2016) at the University of Washington, guides the user through the recipe for a summer roll. According to the video, the application seems to use voice commands to continue to the next step, showing text and videos for steps. On a cutting board a circle is displayed, showing where to lay a rice pancake. The next step displays small rectangles on top of each other within the circle, indicating where to put shrimps, tofu and salad.

#### **CookAR**

For the other capstone project CookAR from State University of New Jersey, Orsini et al. (2017) use image recognition for identifying ingredients and making recipe suggestions. Once the user selects a recipe it will show text instructions, enhanced by some text, videos and animation. It uses voice commands for interaction as well. No information is found on which recipes are supported by the application.

---

<sup>1</sup>Demonstration that the idea is feasible

<sup>2</sup>Assignment at the end of an educational program to demonstrate all gained knowledge and skills throughout the courses



### Sous-Chef

Sous-Chef is an application that was created during a Hackathon by Lock et al. (2016). This application shows a panel containing different recipes. Once a recipe gets selected, it will open a different panel with the whole recipe. Some recipes also contain video links, which open in a different panel.

To summarize, these projects are all POCs that show that it is feasible to use a HoloLens in the kitchen context. However, these projects don't investigate which problems will emerge with such a device during cooking nor how an application is perceived by users. Hence, adequate testing needs to be conducted to make a reliable statement.

## 3.2 Guided Assistance Through a HoloLens

Plenty of research deals with HMD devices in various contexts but none deal with the kitchen environment.

Blattgerste et al. (2017) compared four different instruction methods: a Microsoft HoloLens, an Epson Moverio smart glass, a smartphone, and paper. Participants, in this case students, had to follow different assembly tasks with Lego Duplo bricks. Results of this study showed that tasks were completed the fastest by paper instructions, followed by the HoloLens instructions, except in the task of finding the correct position to assemble the bricks, where the HoloLens was third. The least number of errors overall were with the HoloLens device.

Another paper illustrates a requirement analysis for an HMD training application for workers at an assembly line (Werrlich et al. (2017)). A 3D model of an engine is displayed with two buttons and text instructions. Through gaze and air tap gesture ten participants were able to follow three predefined work instructions. The researchers found out three things: that employees like to have some short animations/movies; that gesture control needs more training; and that FOV is small and the device is heavy.

In the context of this thesis, the previously mentioned studies gave initial insights as to how a HoloLens device is generally perceived and making note of the differences between paper and HoloLens instructions. The papers show that following instructions through a HoloLens device seems to lead to less errors. However, these studies don't indicate how the device can instruct in a kitchen context or what its limitations are in said setting. It doesn't also say how people feel after wearing the glasses for an extended period.

## 3.3 Cooking Instructions in the Kitchen

One way of getting instructions in the kitchen can take place through projections on the kitchen counter, enabling content to be displayed next to the actual working space.

An older research from Horie et al. (2006) demonstrated a projector application which projects content on a table. The device had two modes: one recognized the user's actions, the other instructed the user with steps. The object recognition was identified through markers. For this research the Japanese dish "Okonomi-Yaki" was cooked. Small steps were displayed through

videos and text. Arrows indicated where the user needed to put the ingredient. Once objects were touched, the system gave immediate feedback to the user regarding what to do.

Sato et al. (2013) presented a projector-based application as well, called MimiCook. It displayed a step-by-step guide which directly projects on ingredients and utensils while participants are following a muffin recipe. A scaling device is connected to the application additionally in order to give immediate feedback. Through the scaling device, the system recognizes if a step has been accomplished. According to the study, participants were able to follow the recipe without any error even though the system was not explained to the user.

Another aspect is examined by the research of Buykx and Petrie (2012). They evaluated how recipes could be better illustrated, both with showing intermediate goals and a dependency graph. In an experimental setup participants cooked three different levels of recipes, with three different formats. The first format was an original one from a cookbook. For the second method, the instructions were divided into small steps. Every step was one task, enhanced with pictures. The final format contained a dependency graph of the recipe, with sub goals. Each step showed a picture as well as the quantity of ingredients. Results indicated that multimedia representations seemed to improve the cooking procedure overall.

Another interesting concept is the *Concept kitchen 2025* (2015) by IKEA. It presents an interactive table. Ingredients and objects which are put on the table are recognized and enhanced with additional information like video or text that serves as a guide during cooking, telling where to cut or how much is needed for example. That kind of information would be interesting to display through an HMD device. To gather more information about what users desire or how they like to interact in the kitchen, the smartKitchen research project was consulted.

These previous studies show that small steps coupled with multimedia enhancements can improve the guidance through a recipe. What is still not known is how the experience would be with an HMD device. Moreover, the studies didn't investigate if there is a benefit for users cooking something unfamiliar. It only shows how an application can potentially support.

### 3.4 SmartKitchen Research Project

As mentioned in Chapter 2, the smartKitchen project focuses on interactive displays in the cooking area, with the aim of improving the cooking experience. Scheible et al. (2016) carried out different approaches to gather data, following the HCD process.

In the *smartKitchen project* (n.d.) explorative analysis through observations while cooking were used to identify opportunity areas. Experienced online interviews were conducted to form categories. Both approaches were combined to cover a broad area of user experience aspects. Based on these findings the researchers created a wide range of ideas. These ideas got rated and narrowed down to a few concepts for further evaluation. By doing this, they further prioritized ideas and decided which ones should be prototyped.

Analysing these gathered data, the data revealed some needs during cooking with which an HMD application can help (smartKitchen (n.d.)). A few prototypes were constructed for the smartKitchen project but none of which answers the given research questions of this thesis.

### 3.5 Conclusion

So far no research has investigated an HMD device in a kitchen context. However, some existing POCs projects show that this is feasible. Researchers have investigated that a HoloLens device can be used for instructions. Different display possibilities were tested by some researchers in the kitchen context such as projections on the kitchen surface to assist during cooking. By analysing these papers and results, some ideas were for a possible application emerged.

In the chapter that follows, the insights from this chapter will be used to elaborate the requirements for a HoloLens application, suggesting features which can be used in the kitchen and how it will be implemented.

---

# Application

The last chapter provided information what kind of knowledge and what kind of similar projects exist presently to create a foundation for this research. After gathering the facts in Chapter 3, the kind of features the HoloLens application should have must be consolidated before it gets implemented.

How an application could assist a user following a recipe also needs to be understood. The usefulness of these thoughts are to be evaluated in a detailed assessment. Lastly, an overview is given on how the application got developed.

## 4.1 Requirements

This section consists of four subsections. First, the interaction within the application will be discussed, pointing out why a character was chosen. The second subsection will address qualities of the character to ensure it will be perceived as something helpful. Third, while following a recipe, the application will be supplemented with media content. Lastly, a possible solution for suggesting different recipes to the user based on ingredients will be introduced.

### 4.1.1 Interaction Within the Application

Before examining possible content for the application, the question came up at to how the user will interact with the application. There are several ways which can be considered even though the HoloLens is a hands-free device.

Users can interact through gaze, gestures and voice. Just gazing at an object can trigger an action whenever the user looks at something, be it aware or unaware, which can be a disadvantage. Gestures have the effect that at least one free hand is needed for triggering an action. In the kitchen, hands are most of the time busy with doing something. On the other hand, voice commands can always be given. The idea formed that during the cooking process the interaction should mainly focus on voice. Actions can be triggered either by just speaking or speaking to a virtual object. Devices which listen for a specific keyword are called smart speakers. Examples are Amazon Echo<sup>1</sup>, Google Home<sup>2</sup> or Apple HomePod<sup>3</sup>. The device only reacts when the keyword is spoken. The disadvantage is that the keyword needs to be said in every sentence. The more convenient

---

<sup>1</sup><https://www.amazon.de/Amazon-Generation-Intelligenter-Lautsprecher-Schwarz/dp/B01DFKBG54>

<sup>2</sup>[https://store.google.com/product/google\\_home](https://store.google.com/product/google_home)

<sup>3</sup><https://www.apple.com/homepod/>

solution is not to say a keyword at all. Then the device listens all the time and tries to analyse the content. However, this could raise privacy concerns for some people since every single word is analysed in order to understand what the user is intending to do.

For these reasons a character was selected as a contact point for every activity. A virtual object, in the form of a character or something similar will be displayed to the user. Once the user gazes at the virtual object and speaks, an activity is triggered. Additionally, gesture control will be possible for tasks which don't interfere with the actual flow of cooking to supplement interactions.

### 4.1.2 Character Characteristics

Not only the appearance of the character needs to be thoroughly considered but also how the user would benefit from a character.



Figure 4.1: Microsoft Clippy character

In 1996, Microsoft introduced an animated character integrated within Microsoft Office, called "Clippy" (see figure 4.1) (Meyer (2015)).

It was supposed to help the user with their tasks. However, users hated it because it was distracting from their actual work (Rossen (2017)). According to a Microsoft employee, Pratley (2004), Clippy was originally considered for users who use Microsoft Office for the first time. The problem was that Clippy always forced information

to the user at the wrong time.

For the character in the HoloLens application to be seen as something helpful, it should only react to the user's requests. If the character was proactive, it would need contextual information from the user. Otherwise it can't react meaningful to what the user is trying to achieve. To keep the implementation simple the character will only answer when a user is talking to it.

The user should also feel some relationship to the character while using it. The character should call the user by name or engage in some small talk to achieve this. Moreover, any kind of character sentiments would encourage the user to see the character more as a human than as an artificial object.

In order to make small talk, the character needs to communicate to the user through voice. What is spoken should also be displayed as text in case the user doesn't understand the character and wants to read the message again. If there is no voice and only text is displayed, the user can't focus on the task since the user needs to read the text. Additionally it's important that the time for the character to reply takes only a few seconds. If it takes too long the user will probably talk to the character again while the response is being received, which will cause some errors. Or the user gets annoyed because it always takes so long.

The character doesn't have to be a human. The scope of this thesis is not the application itself. The aim is to understand the overall perception of the HoloLens device in combination with an application. For this reason, the character can be as basic as possible. Moreover, a human character is complex to animate, be it body movements, facial expressions or gestures. If the virtual human only functions like a half human then the user won't be engaged while using the application because the interaction with the virtual human will not feel real.

### 4.1.3 Cooking Guidance

Considering the application is only used in the kitchen, it should assist in following a recipe. Buykx and Petrie (2012) found the best way to support during cooking is in small steps, preferably with a dependency graph enhanced by pictures. With that in mind, the goal was to implement such a dependency graph with sub goals. This graph should be based on buttons which are connected to each other, situated somewhere in front of the user's view. The buttons would make it easier to jump from one step to one several steps away.

It makes cooking easier if the quantity of ingredients is displayed either in the current step or on a separate list along with a video, rather than an image, displaying what the user has to do (Buykx and Petrie (2012)). A video is preferred because it can give deeper understanding how a step should be executed.

The recipe should be something unfamiliar. Using a HoloLens device for a dish which is familiar won't be much help because people would already know what needed to be done. But the recipe shouldn't be too complicated to cook either. With average cooking skills the user should be able to achieve it. If the recipe is too easy to make, the HoloLens will be redundant.

Because this recipe will be tested with a number of participants, it should be also possible to in a reasonable amount of time. It should also come along with videos, showing every single step. The videos will be used to guide a user through the recipe.

### 4.1.4 Recipe Recommendation

Pagenkopf et al. (n.d.) and smartKitchen (n.d.) gives insights from the conducted user research in the smartKitchen project. The results showed that some participants mentioned that it would be useful if the application could suggest recipes based on available ingredients. As a matter of fact, the HoloLens device is capable of taking pictures and videos. This provides the possibility of analysing objects which are within sight of the device. With this in mind, the HoloLens application could scan vegetables in front of the device and based on that information, it can suggest a selection of recipes. The application should be able to recognize at least two different vegetables to make a meaningful suggestion. Additionally, the whole recognition process should take only a few seconds, otherwise the user experience will be annoying because it takes too long.

Several other features came in mind which could be added in this application. It was decided to leave them out of scope since the time was limited for the implementation. The existing features should already be able to cover the basic functionality for following a recipe.

In brief, the application should contain the following items:

- A character, which doesn't have to be human.
- Interaction with the user mainly through voice.
- Capacity to scan vegetables in order to make a recipe suggestion.
- Capability to guide through a recipe with small steps, enhanced with text and video.

## 4.2 Design and Technology Choices

Now that there's a basic idea for an application, the details of how to such an application should be developed can be thought of in more detail. One place to start is the kind of character that should be used. Any character is possible, but what would makes sense in the kitchen context? Another question is what recipe should be cooked for this research. Once these decisions have been made, the ingredients needed for image recognition can be identified. It is also important to decide which language the character will speak and understand, as well as which tools are needed for implementation.

### 4.2.1 Character

The idea of this application is that the user is interacting with a virtual character. Therefore, what the character looks like is very important.

#### Appearance

First of all, the character needs a visual representation. Since no previous experience existed with 3D modelling, free 3D models were looked up. Models with a rigged body were searched for so that the 3D model can be animated. The choice fell on a penguin<sup>4</sup> which seemed not too complicated for animating. It doesn't have many body parts and the mouth can be moved as well. The penguin also looked friendly enough for interaction.

After giving the character some basic walking animation in Blender<sup>5</sup> the question came up as to how the penguin will walk around on top of a kitchen table. Ideally, if there are many objects, utensils or ingredients placed on the table, the penguin needs to be aware of them so that it doesn't walk into these objects. The HoloLens actually has spatial awareness and can detect objects in the surroundings but implementing it would take more time. However, the aim was not to create a perfect character but about how the character could support while cooking. So the character was designed to fly around in the air since it is easier to implement when there is no need to interact with its surroundings, assuming that it also stays in one position.

The search for a flying object ended up at a website by Nick Landry. He created a sample project called HoloBot, which is a "floating robot head using speech recognition" Landry (2016). The source code is available for free, following MIT licence conditions. This project seemed as a suitable base for the application. Although the character doesn't move its mouth, the whole object is constantly moving a bit up and down, making some slight background noise while giving feedback to the user.

#### Interaction

The interaction between user and character can be divided into three different parts:

1. The user speaks to the character - Voice To Text,
2. the character understands what the user is saying - Language Understanding; and

---

<sup>4</sup><https://www.blendswap.com/blends/view/68089>

<sup>5</sup>Tool for 3D Modelling - <https://www.blender.org/>

3. the character speaks to the user - Text To Speech.

The HoloBot project already comes along with understanding voice commands utilizing different services from Microsoft. When the user gazes at the character, a short beep sound is played, indicating that the character is listening until the user pauses from talking. This recorded text is sent to the cloud, gets analysed, and returns a corresponding action back to the application. The received text will then be spoken by the character to the user.

The HoloBot project uses cloud services from Microsoft, Microsoft Bot Framework, and Microsoft Language Understanding Intelligent Service (LUIS) to understand commands. A so called chatbot service is receiving through text from the HoloLens application. This service enables, on a textual level, a natural conversation between user and a computer (Microsoft (n.d.)). To be able to understand the human input, the text needs to be translated so that the computer can understand it. A number of different services exist for extracting these information from a sentence, the so-called natural language processing (NLP). Some common services from bigger companies are Wit.ai<sup>6</sup> from Facebook, Dialogflow<sup>7</sup> from Google, Lex<sup>8</sup> from Amazon, LUIS<sup>9</sup> from Microsoft and Watson Assistant<sup>10</sup> from IBM. A shallow comparison of this different engines showed that they are more or less the same, depending on requirements. For this thesis it was chosen to use Microsoft Bot Framework in combination with LUIS since it was already used by the HoloBot project, and since the aim was not to investigate differences between different engines.

#### 4.2.2 Recipe

At the time of searching for a recipe that is unusual yet easy to make, a "semi guided cooking app" was done as part of the research project smartKitchen by Lisa Schuhmacher. The aim of this application is to add value to the cooking process through individualization and interactivity with the user and kitchen (Schuhmacher (2018)).

The cooking video therefore is split into small cooking segments, which enables the application to adjust to the user's speed. A non-linear cooking experience is achieved by connecting to an IOT-Platform, communicating with different features of Alexa and other kitchen devices. This leaves it up to the user how to interact: either using the application with gestures or voice.

The produced videos contained steps for making a sushi burger. Every step is a single video, supplemented with text for a smart speaker to say what has to be done during each step. Since the recipe of the project fulfilled the needs of this thesis, it had been selected. The whole steps for the recipe can be found in Appendix A, and the end result of the sushi burger is seen in the figure 4.2.

---

<sup>6</sup><https://wit.ai/>

<sup>7</sup><https://dialogflow.com/>

<sup>8</sup><https://aws.amazon.com/lex/>

<sup>9</sup><https://www.luis.ai/>

<sup>10</sup><https://www.ibm.com/watson/>





Figure 4.2: The final sushi burger

### 4.2.3 Image Recognition

Image recognition is needed in this research for detecting vegetables, giving the user feedback on what can be cooked based on the found ingredients. The idea is that the HoloLens either takes a picture or takes a video of what it is currently seeing and then the content is getting analysed in the cloud. On top of this, an algorithm will return the probability of what has been recognized, best case of which is identifying what kind of vegetable is in front of it.

The following table, table 4.1, dated in the first Quarter of 2018, indicates on a high level what kind of services the three key players are offering for image analysing. In terms of object recognition, the three companies can mostly determine what kind of object it is or recognize the kind of scene.

	Amazon Recognition	Google Vision	Microsoft Cognitive Services
Object detection	x	x	x
Scene detection	x	x	x
Activity detection	x		
Facial recognition	x	x	x
Facial analysis	x		x
Person tracking	x		
Unsafe content detection	x	x	x
Celebrity recognition	x		x
Text in images	x	x	x
Logo detection		x	
Landmark detection		x	x
Image attributes (Colour / Crop hints)		x	x
Web detection		x	
Create thumbnail			x

Table 4.1: Comparison of different object recognition services

The three services were roughly tested with a few vegetables. As of the time of writing, the results were not satisfying. A tomato and cucumber was not always recognized. Amazon Recognition<sup>11</sup> and Google Vision<sup>12</sup> were tested by uploading some test images. A sample iOS application<sup>13</sup> existed to test the Microsoft Cognitive Services<sup>14</sup>. It also had difficulties recognizing these objects. Vegetables were apparently not the focus of these libraries.

For this reason the other possibility was to train an own model which could recognize vegetables. Some open source frameworks, like TensorFlow<sup>15</sup>, provide machine learning. With such a powerful framework, it's possible to build an own image recognition system. For a system to be capable of labelling a picture, thousands and thousands of pictures are needed per category. The more pictures are provided, the more accurate the label will be. Creating a new image dataset is time and CPU consuming. Not only does it involve finding a lot of images for a category, the objects in the images also need to be manually annotated. Once done, it would be possible to train the system with these image models and the computer will learn by its own.

These steps are also being offered by a cloud service from Microsoft called Custom Vision. The service expects a set of images which can be uploaded and then categorized (Microsoft (n.d.)). The annotation is done by some backend logic. A simple click triggers the training of your model and it takes only a few minutes. The downside of the free version is that the number of pictures for uploading (in total 1,000 pictures) and API calls (1,000 per day) are limited. With such a limited number of pictures, only a few categories should be created to guarantee a higher probability of object recognition. For this reason it has been decided that the model will be trained using only three vegetables. Otherwise it would be hard for the system to differentiate between the vegetables. These selected objects were also easy to differentiate. For this thesis an avocado, a cucumber and a tomato were chosen.

#### 4.2.4 Language

The language for the interaction with the character was set to English. It could have been German but the bottleneck was the Voice to Text and Text To Speech feature. At the time of writing, the HoloLens device only supported English for speaking. The Microsoft Bot Framework and the LUIS service are capable of understanding German.

#### 4.2.5 Tools

For the HoloLens development, two different programs were needed, Unity<sup>16</sup> and Microsoft Visual Studio Community 2017<sup>17</sup>. Unity is a development platform for games. This tool is recommended by Microsoft for developing 3D applications for the HoloLens. The version 2017.2.1p2 was used for this thesis. The HoloBot sample project is provided as a Unity project. In Unity all necessary 3D modelling can be done, as well as adding/modifying code which opens Visual Studio.

---

<sup>11</sup><https://aws.amazon.com/rekognition/>

<sup>12</sup><https://cloud.google.com/vision/>

<sup>13</sup><https://github.com/DanilaVladi/Microsoft-Cognitive-Services-Swift-SDK>

<sup>14</sup><https://azure.microsoft.com/en-us/services/cognitive-services/>

<sup>15</sup><https://www.tensorflow.org/>

<sup>16</sup><https://unity3d.com/>

<sup>17</sup><https://www.visualstudio.com/de/vs/whatsnew/>

When building the project in Unity with HoloLens settings, a Visual Studio solution is created. In the project file, further code adjustments can be made. From there the application is deployed to the HoloLens device or the emulator.

The HoloBot project is using a Mixed Reality Toolkit<sup>18</sup> containing a number of scripts which are commonly used. In the sample project a older version is used, which has been replaced with version HoloToolkit-Unity-2017.2.1.3 .

The Microsoft services, Bot Framework and Custom Vision, can be tested independently from the frontend application. For the Bot Framework a "Bot Framework Emulator" desktop application exists, allowing us to have a conversation through text messages in the form of a chat with the bot. In contrast the image recognition service can only be tested by uploading an image or a URL link to an image.

To summarize the choices:

- A flying robot head will be used as a character.
- The user will talk to the character and it will reply through voice, for which the Microsoft Bot Framework and Microsoft LUIS will be used.
- The application will guide through the steps of a sushi burger.
- An avocado, a cucumber and a tomato will be recognized for image recognition, using Microsoft Custom Vision.
- The character will speak English.
- Unity 2017.2.1p2, Visual Studio Community 2017 and HoloToolkit-Unity-2017.2.1.3 will be used as development tools.

### 4.3 System Overview

In figure 4.3, the overall architecture of the application can be seen. The application is a Unity project which uses different classes from Unity for main functionalities like saving videos in the local file system, recording of the spoken text, and reading out the text as example.

Regarding speaking, the application converts the voice input from the user to text, then sends it to the Microsoft Bot Framework service through a JSON object. The exchange between Microsoft Bot Framework and LUIS is seamless. The Bot Framework is a separate Visual Studio project, which forwards the text to LUIS in order to analyse the content of the sentence. LUIS has to be trained with words and phrases which the application should react to. Similar sentences are united under an intention.

Once the sentence is analysed, LUIS tries to assign an intention to it. If an intention is found, this will trigger the corresponding action in the Microsoft Bot Framework. The action contains information what will be said to the user through the application character. This action is sent back to the application through a JSON object.

---

<sup>18</sup><https://github.com/Microsoft/MixedRealityToolkit-Unity>

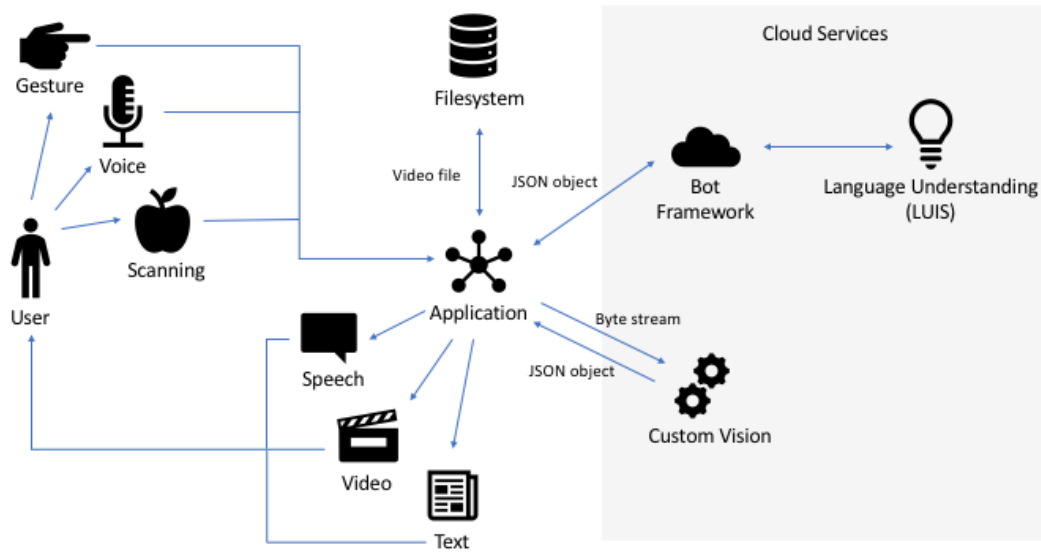


Figure 4.3: System overview of the application

For recognizing ingredients, an image is taken with the HoloLens device and then sent through a byte stream to the Custom Vision service. There the picture gets analysed and the service responds with a JSON object telling what the image contains and the probability of what it contains.

The application uses internal classes to convert text to speech in order to talk to the user. Additionally, video and text functionality from core frameworks are used to display content to the user.

## 4.4 Implementation

This final section is devoted to the implementation of the application and services, listing on a high level all main characteristics of the application and its services.

### 4.4.1 Backend Services

First, an account is needed for using Microsoft Azure<sup>19</sup>. Microsoft Azure is a platform containing a whole range of cloud services from Microsoft. For this thesis a new resource called "Web App Bot" with LUIS was created, it was named "smartKitchenBot." Once this was done a sample bot project was automatically generated which can be downloaded as a zip file under Build, see figure 4.4. This file contains all code to run the sample bot application using this service. The file includes information which is needed to publish it back to server within Visual Studio through the Publish settings.

<sup>19</sup><https://azure.microsoft.com/en-us/>

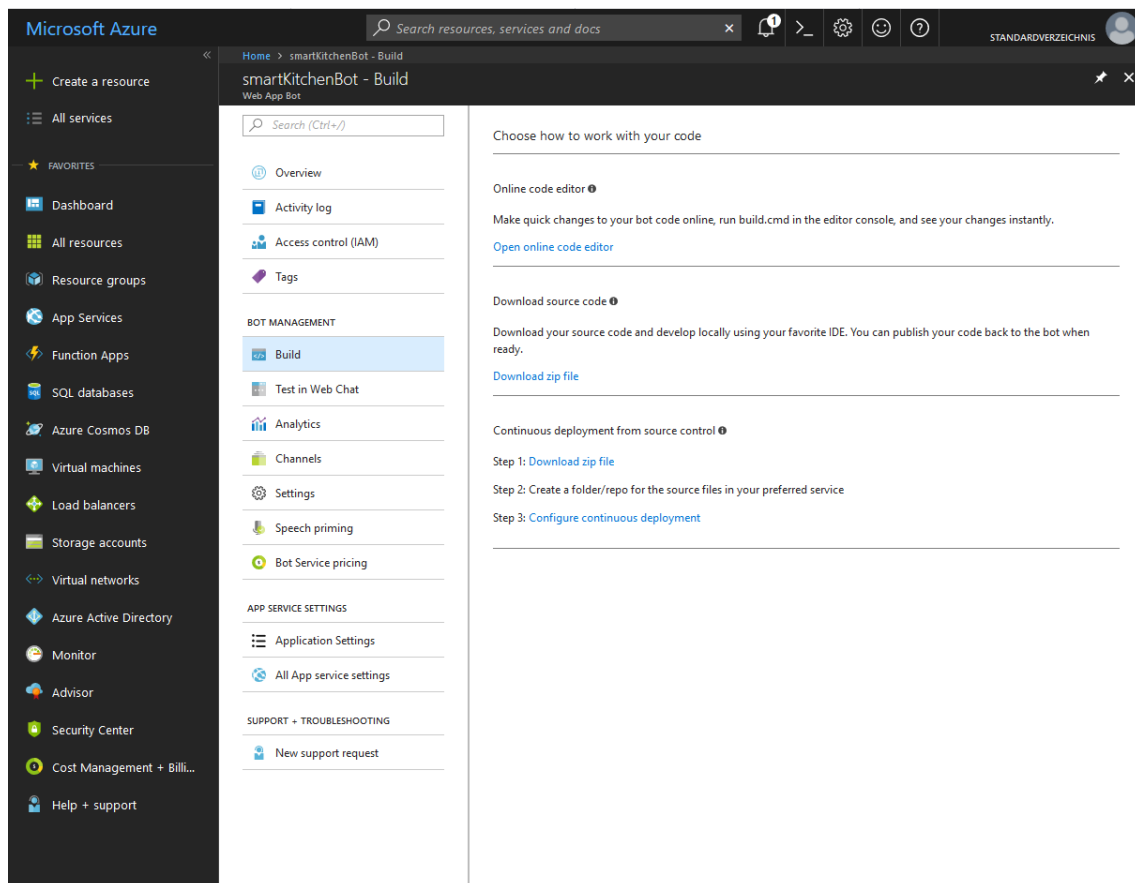


Figure 4.4: Microsoft Bot Framework sources in Microsoft Azure

Before the actual coding of the application took place, how a conversation flow between a user and a character can run was thought about. The conversation sequence from the "semi guided cooking video" project from smartKitchen was used as a basis with some modifications by the researcher. Some additional small talk were added before and after the cooking and modifications to the cooking flow were made. The application starts with a short small talk between user and character before the user gets asked what to cook. In case the user doesn't have an idea of what to cook the application will recommend different recipes based on ingredients. This process will be called inspiration flow. Once selected, in this case the sushi burger, the application will guide through each step of the recipe, called cooking flow. In Appendix A an example of a whole conversation flow can be found.

### Microsoft Language Understanding Service

For the Bot Framework to know how to react to a request from the application, the framework needs to understand what the user intends to do. For this reason LUIS is used. LUIS will analyse the text which is received from the Bot Framework and return an intent for what it understood. Seventeen intents were created for this thesis, consisting of intents for different states for example yes, no, good, done, bad and others. There are intents for starting the inspiration or cooking process. One intent is for starting the sushi burger, and another one is for when the user selects some other recipe suggestion.

For example, the user wants to start cooking in the application, the intent "Smartkitchen.Cooking" needs to be triggered. Within this intent all possible words and sentences which were thought of, how a user would say that cooking should start are keyed in (see figure 4.5).

The screenshot shows the Microsoft LUIS interface for the 'SmartKitchen.Cooking' intent. The top navigation bar includes 'Language Understanding', 'My apps', 'Docs', 'Pricing', 'Support', 'About', and the user name 'Adrian Föhl'. The main navigation menu on the left includes 'App Assets', 'Intents', 'Entities', and 'Improve app performance'. The main content area shows the 'SmartKitchen.Cooking' intent with a 'Delete Intent' button. A text input field contains the placeholder text 'Type about 5 examples of what a user might say and hit Enter'. Below the input field are 'Entity filters' and 'Show All' options. A table lists several utterances, each with a 'Labeled intent' dropdown menu set to 'SmartKitc...'. The table also includes a 'Name' column and a 'Labeled utterances' column. At the bottom, there is a section for 'Entities used in this intent' which currently shows 'There are no entities in use.'

Figure 4.5: Microsoft LUIS with an example intent

Once words and sentences are entered into the LUIS system, the model needs to be trained and published. Afterwards, if the Bot Framework sends a text to LUIS which looks similar to cooking in this case, the intent "SmartKitchen.Cooking" will be returned. And the Bot Framework will handle all necessary steps for this intent which is explained in detail next.

## Microsoft Bot Framework

Although the language is set to English for the interaction, it might be the case that at some point in future the application will use a different language. For this reason, the Bot Framework project was setup so that it is working with localized files. To add another language, a new Resources.resx file with the new language extension simply needs to be added. The correspondent resources file needs to have the key, which text should be replaced, and a value, the text in the new language. For using these changes, the Global.asax file needs to be modified using the commented lines currently in the application.

The code was structured into three different dialog parts: RootDialog, InspirationDialog and CookingDialog. A dialog class contains all information to manage a conversation flow.

### **RootDialog**

This is the entry point of the application. Depending on the returned intent from LUIS, it either initiates the greeting, inspiration or cooking part.

### **InspirationDialog**

This class contains the flow for giving the user a recipe suggestion. Overall questions are asked in order to simulate the feeling that the different recipes are suggested based on the user's input. The application will leave this class once the sushi burger is selected and returns to the RootDialog, then triggering CookingDialog immediately.

### **CookingDialog**

This dialogue will guide through every single step of the recipe until it is finished. Once finished it will return to the RootDialog, where a termination of the application is triggered.

Several cooking steps are combined to a section: Preparation before the cooking, rice cooking, making filling, adding flavour, and building the burger. The reason behind this is that buttons will be displayed in the application which enables the user to jump to the beginning of a section. Originally, it was planned that every single step will be shown as a button. This plan changed due to a problem which will be described next.

A normal conversation has a lot of different possibilities on how it can be carried out. Except for the small talk part, the questions for this thesis are following a defined flow. The user gets asked questions to collect information about cooking preferences or moving to the next step. Different answers are possible but the conversation flow will always be the same. The Bot Framework offers a feature called FormFlow that gives the possibility of easily managing a guided conversation (Standefer et al. (2017)). The inspiration FormFlow for example is asking about how much time the user has, how difficult it should be to make, for how many people, and if the user has any special requests. This data is defined in a form class as individual fields. This sequence is now processing every field one after another. Between these fields the application always waits for the answer from the user before going to the next step. Apparently, it is not possible to jump within a FormFlow, for example between step 2 and step 4. It can only go back to the previous step. For this reason, the whole recipe was divided into several sections, each section consisting of a FormFlow. With this approach, it is possible to go back to the beginning of each section. For example, when making a burger and the user wants to go back to a flavour step, the flavour section button needs to be selected. But then the user needs to go through every step until the desired step.

### **Microsoft Custom Vision Service**

Before any object can be recognized, a set of vegetable pictures is needed. The website ImageNet<sup>20</sup> offers a large set of image categories which can be downloaded. The free version of the Custom Vision service allows only a maximum of 1,000 pictures to be uploaded. So that each category has the same training set amount, the number of images for each category (cucumber, avocado and tomato) was set to 300 images. After downloading and collecting the amount of images for each

---

<sup>20</sup><http://www.image-net.org/>

category from ImageNet a new project was created in Custom Vision. The three picture sets were then uploaded. A tag was assigned to all pictures and then the model was trained, see figure 4.6.

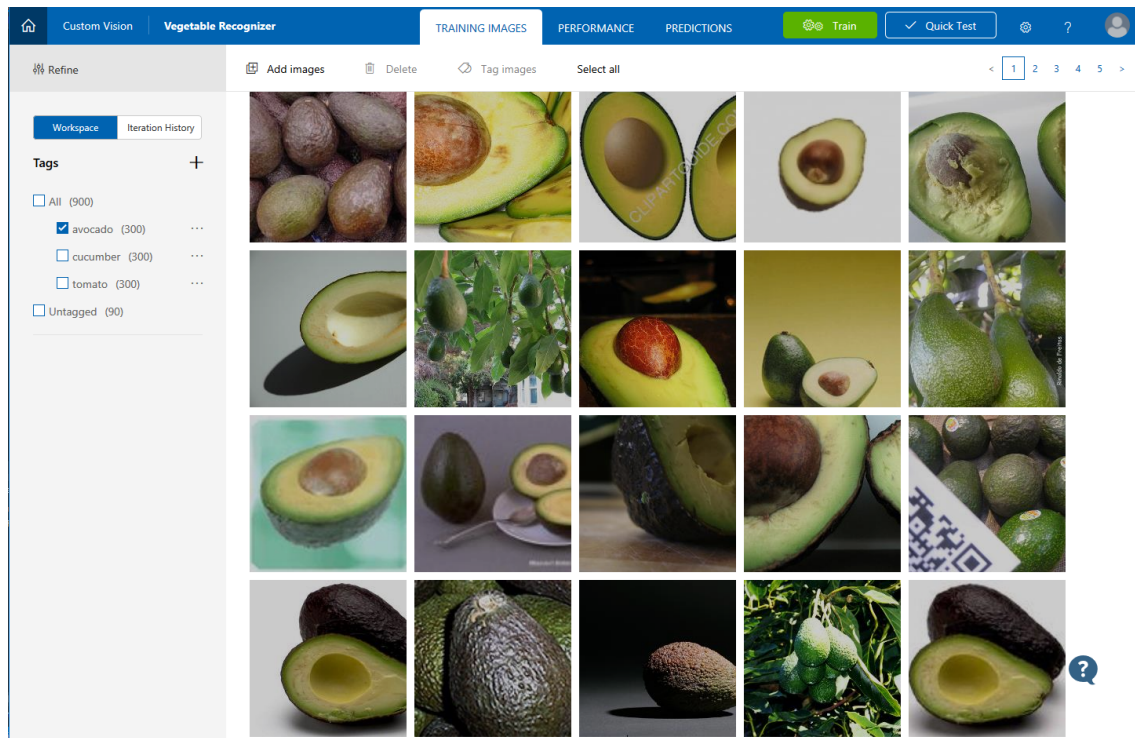


Figure 4.6: Training set for avocados in Microsoft Custom Vision

#### 4.4.2 Frontend

The character itself didn't need any further modification. Additional objects for text and video were added which can be found in the "Scripts" folder of the Unity project.

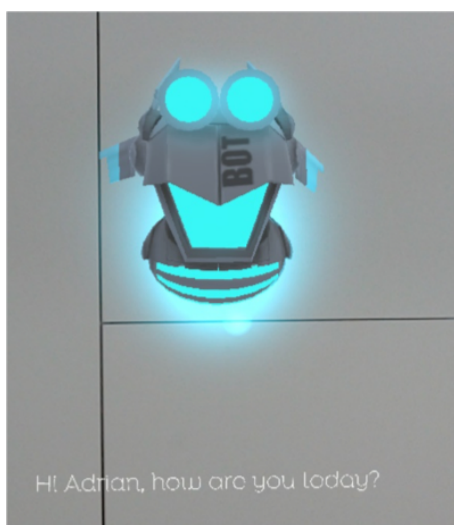


Figure 4.7: Character of the application

Once the application starts, the character hovers a little above eye sight around 1m in front (see figure 4.7). It was decided that the character stays at that position the whole time and won't follow the user around.

The text below the character is showing either what got understood from the user or what the character is saying. The text is displayed the whole time until some new text comes. It will always stay at that position below the character. Throughout the usage of the application the character and the text are visible.

A constant, not so loud sound is playing in the background, simulating some kind of machine hovering noise. Whenever the head of the user is gazing at the character a short beep sound is played, giving feedback to the user that the application is recording what it is hearing. The recording stops when the user pauses talk-

ing or if nothing is said after a few seconds.



So that the user knows where his view is currently focused at, a cursor is always visible. A small light purple sphere is used when the user is looking around. It turns into a light blue sphere when it hits the character. If a button is gazed at, the cursor transforms into a 2D blue circle with a hole in the middle.

The application starts with greeting the user by name and asking how the user feels. Depending on how the user answers it will reply and ask if the user is hungry and if the user knows what to cook today. If the user answers yes then the cooking flow will start, otherwise the inspiration flow begins.

### Inspiration Flow

Several questions are asked within this flow like how much time the user has if the user wants to cook something easy or difficult, or if any special requests are needed. Lastly, the character will ask the user to place some food on the working space. A light grey transparent panel will appear telling the user that it is scanning. The panel is shown where the user is looking at and it takes nearly the whole space in the display. It will also tag along with the view of the user, meaning wherever the user looks or moves the panel will follow. It will be visible until the application successfully recognizes at least two vegetables.

Once the panel is displayed, it will trigger in the background the `StartAnalyzing` method in the `ObjectRecognizer` class. This method takes a picture every three seconds of what the user is seeing. Three seconds was chosen since the Custom Vision service only allows a maximum upload of 1,000 pictures per day. Having a couple of tests in one day might exceed this limit.

The application uploads the image to the Custom Vision service and then waits for a JSON response. This response contains what kind of objects were found and the probability of it. When the probability is higher than a defined value in the frontend application and two objects were found then the scanning panel will disappear. A menu with three different recipe options will be displayed (see figure 4.8).

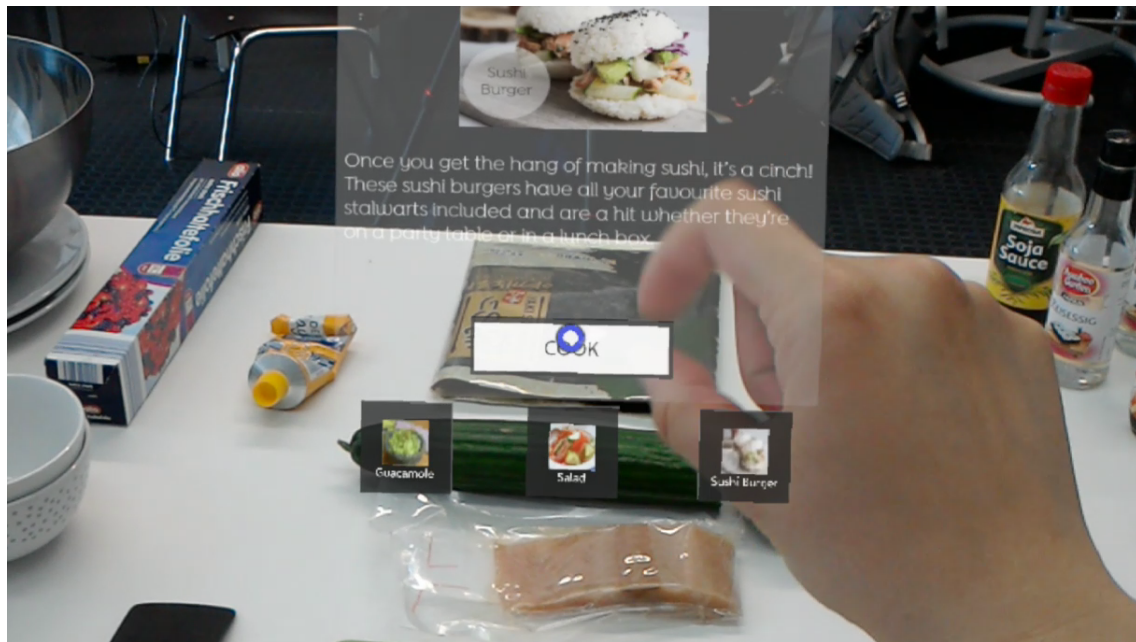


Figure 4.8: Recipe selection in the application

Possible choices are a guacamole dip, a tomato cucumber salad, and a sushi burger. Whereas the salad is displayed first. The panel shows a picture of the dish with a short description. For the sushi burger there was an additional button where the user can start the cooking process. Like the scanning panel, the menu always follows the user wherever the user looks. Once the button is selected, the cooking flow will start. Instead of selecting any button, the cooking flow can also be started through voice by saying something with "sushi burger."

### Cooking Flow

When the cooking process starts, the user is given information that in order to continue either "OK" or "next" needs to be said. Before each cooking step is begun, the user is asked if all tools and ingredients are available. When asked for tools, an additional text on the right side of the user view will appear. Listing all needed tools one below the other. The next step shows all ingredients and the quantity. In addition, a video is shown, on the left side from the list. Which supplementary shows a hand putting all ingredients on a kitchen table.

Both video and text are situated in front of the character. The video is taking the middle position of the view and the text is situated on the right (see figure 4.9). Both contents will follow the user when moving around. In case the user looks from the side, the text and video will move around the y-axis so that it can always be read and seen by the user. Originally, the video and text were supposed to be always around the same position in the view, but it didn't always follow the view of the user for some reason.



Figure 4.9: A cooking step in the application

While the cooking steps were shown, four buttons were presented, as seen in the figure 4.9. This enables the user to jump between steps, as mentioned in the section Microsoft Bot Framework, 4.4.1. When a button is selected, two events are fired one after another. The first event cancels the current step and the second one tells the Bot Framework which cooking section should start next.

The text on the top right displays all ingredients which are needed for this process and the quantity. Below the ingredients text is a summary of steps that the character was saying as an overview. However, the text below the character always shows the whole description as well.

The JSON message which is received from the Bot Framework contains the text which should be spoken by the character. Therefore it is parsed and compared with keywords which are saved in the application. Depending on the received message it will trigger different actions. Ingredients information and step summary is obtained from the application. These information are separately added to the screen in order to see it at a quick glance. The spoken text from the character is not suitable for that.

After the user finished all cooking steps, the character says that it wants to eat it and that it looks really good, praising the user by name. Since the cooking was tiring the character will say that it will take a nap. The application will then terminate itself after a few seconds.

### 4.4.3 Challenges

During development of the application some noteworthy issues came up which will be listed in a short summary.

#### Deployment to Device

From Unity to the HoloLens it takes two steps. First, a Visual Solution project file needs to be created in Unity. This build can take around two minutes to create. Second, the created project file gets opened in Visual Studio. From there it can be deployed to the device which takes slightly

more time. It was developed on Acer Aspire V15 Nitro Laptop, with core i7 2.6 GHz and 16GB RAM. Deploying to the simulator takes a little less time, but it wasn't possible to interact through voice.

### **Dizziness While Testing**

While testing the application on the device some kind of dizziness occurred to the researcher due to some shaking from the character or from the text when moved around. In a specific angle the content is shivering and can give some headache after some time. Why it is shivering couldn't be clarified.

### **Dirty Sensors**

It happened that after some time the air tap gesture didn't work properly anymore. It was reacting but several attempts were needed. One reason could be that the sensor on the HoloLens device is dirty and needed to be cleaned. Cleaning the device is easily remembered when developing a mobile application because it is seen while testing on a smartphone but not so obvious on an HMD device.

## **4.5 Conclusion**

This chapter described the development process of the application for this thesis, from the idea to the actual prototype. The application for this thesis built up on a sample project called HoloBot, a flying robot head character, from Nick Landry. This got selected because it was a free 3D model and the preferred choice was some flying character. If something would have to be projected on the table, the character needed to be aware of surrounding objects otherwise it would feel distorted.

The sample HoloBot project was modified in order to inspire for cooking something and guide through every single step of a sushi burger recipe. Two different services from Microsoft were used which were able to understand the spoken language and identify objects in images. The services got trained with a basic set of sentences and images, which is needed to run the application on an acceptable level. After wiring the services and application together, the character was able to guide the user through the cooking with videos and text.

The next chapter will describe how everything was setup to test the application together with the device with the aim of answering the given research questions.

# Method

The last chapter gave details about the application developed for the HoloLens. A character is used as a contact point for interacting through voice. It suggests something to cook and guides through the steps in cooking a sushi burger, supplemented by videos and text.

This chapter will delve on how the application and the HMD device can be evaluated. Two different approaches were utilized for the research design. One is an exploratory design to better understand the advantages and disadvantages for using a HoloLens while following a new recipe. The purpose of this design is to get insights on how a possible application could help, as well as which technical challenges exist with a HoloLens. In contrast there is a conclusive design, which was applied to get quantitative data about the application as well as from the findings regarding the device. This chapter will cover why the designs were selected, and the kind of participants that attended. It will also give information about the qualitative and quantitative methods used to explore the hypothesis and how the study was conducted.

## 5.1 Research Design

A mixed method design was chosen for this thesis to understand why the HoloLens, in combination with an application, is useful and which particular features of the application made it so.

Since no earlier studies exist where a HoloLens was investigated in the cooking context, the approach should give insights and better understanding for the research. This can be achieved by a semi structured open-ended interview and observation of the participants during the cooking process. This was selected to answer the research questions of this thesis.

Another approach would be comparing different groups while cooking: one group cooking with the HoloLens the other one without. Due to the fact that the application was a basic prototype, it couldn't be helped that technical problems which could bias the results would occur. This experimental design approach would emphasize another aspect of research.

A second method is needed to evaluate how participants perceive the application. Which aspects were helpful or would need improvement for an upgrade of the application. The exploratory design is not capable of answering this angle and quantify it. Because of that a conclusive design was selected. This research design verifies and quantifies findings from the exploratory design. By combining both approaches, the validity of the findings will improve.

## 5.2 Instruments

Type	Instrument	When used
Quantitative & Qualitative	Observation	Throughout cooking
Qualitative	Think aloud	Throughout cooking
Quantitative	Semi-structured interview	During salmon cooking and after finishing the burger
Quantitative & Qualitative	2-phase questionnaire, structured and with open-ended questions	After cooking is done
Qualitative	AttrakDiff	After the questionnaire

Table 5.1: Used instruments overview

Table 5.1 demonstrates an overview of different instruments which were used during the whole cooking session. Throughout the test, participants were able to share their thoughts about the application and device, also known as a think aloud method. In case the participant shared insights, notes were made and more questions were asked to further narrow down their thoughts and to gain more background information. A quantitative observation was used as well to analyse subconscious data to understand the occurrences of behaviour. Meanwhile, the qualitative observation aimed to look for new behaviours with the device. The quantitative observations were done by analysing the video files.

While waiting for the salmon to be cooked a semi-structured interview was carried out. These type of questions asked about the applications' object recognition feature and overall character (see Appendix B.1). Directly after finishing a sushi burger, another semi-structured interview took place asking about the overall application and device experience (Appendix B.2). A semi-structured interview was chosen because each participant will have a unique experience which would be hard to find out about in a standard survey.

Once the participants wrapped up their burgers they were asked to fill out a structured questionnaire with open-ended questions. The questions were prepared on a Google Form<sup>1</sup> and shown on a laptop browser. The questionnaire was more favourable over an interview so that the researcher had time to prepare the next cooking setup as the participants answered the form. The cooking sessions were planned with 15-minute breaks in between to offer 25 slots during one week. One week of tests was committed to, since the kitchen was occupied for other project as well.

The questionnaire (see Appendix B.2) was split into different parts:

Part A: Personal data and overall cooking experience (quantitative data)

Part B: Application experience (qualitative data)

Part C: HoloLens experience and cooking of a new recipe (qualitative data)

The personal information and estimation of cooking experience helped to gain information about the participants' background, which will be helpful in drawing conclusions combined with other

<sup>1</sup><https://www.google.com/forms/about/>

data. Part B investigated the perception of the application, and Part C focused on the HoloLens. Additionally, users were asked to report about their last experience when they followed a new recipe and the kind of challenges they faced.

Finally, the user experience of the application was quantitatively evaluated by a tool called AttrakDiff<sup>2</sup>. AttrakDiff is an online tool based on a 28-item questionnaire which measures the product in four different dimensions (Hassenzahl et al. (n.d.)):

- Pragmatic Quality (PQ)  
Usability of a product and if users can successfully achieve their goal.
- Hedonic Quality - Stimulation (HQ-S)  
The ability of this product to improve a user's knowledge and skills.
- Hedonic Quality - Identity (HQ-I)  
The user's identification with a product.
- Attractiveness (ATT)  
Global value of the product.

Different models exist to measure the user experience, like User Experience Questionnaire (UEQ)<sup>3</sup>, but AttrakDiff was chosen because it differentiates between pragmatic quality (usability) and hedonic quality (user experience), both of which are needed to answer the research questions.

Observation was used to collect subconscious data about participants' behaviour regarding application and device. The observations were conducted throughout the cooking process, which took one hour. During the live session special things were noted. A more detailed analysis was done with the videos afterwards.

In the beginning and ending of a session emotional reactions were observed. Notes were taken on the nature of behaviours (interaction, verbal, non-verbal, surprise, laughter, voice tone). In addition, differences between what the users were instructed to do through the character and how they followed it were noted down.

### 5.3 Participants

Select undergraduate and graduate level classes from computer science and media, audiovisual media, media management and corporate communication at Stuttgart Media University were invited through email to participate in this cooking study. Students were the target group for this research because they have less cooking experience than older people and likely need more assistance with following a new recipe. More experienced person needs a different level of guidance.

A total number of 16 students (eight females and eight males) registered for available cooking slots with a length of 90 minutes. From the 16, one male participant did not show up while another male participant faced a technical problem with the character. Instead, he cooked the meal without the HoloLens. However, his participation wasn't accounted into the evaluation. In short, the study was conducted with 14 student participants; eight females and six males.

---

<sup>2</sup><http://www.attrakdiff.de/>

<sup>3</sup><http://www.ueq-online.org/>

The mean age was 24.21 years and ranged between 19 to 29 years (standard deviation (SD) 3.04). All participants were native German speakers who demonstrated at least basic English language skills. This was necessary for the interaction with the character. Lastly, three of the participants have tried the HoloLens before. Participants were able to take away the cooked meal as incentive for taking part in this study.

## 5.4 Procedure

The data collection process took place over a period of five days, from Monday to Friday. In total, 14 observations were made during that week. Five cooking sessions was the maximum per day, with a duration of one and a half hours per session. One hour was reserved for cooking and 30 minutes for the interview. Between two cooking sessions there was a break of 15 minutes in which all utilities were washed and everything was set to its original state.

All tests were done in a kitchen setup at Stuttgart Media University, which was a distraction-free environment. A maximum of four other persons working on their pc were in the room while testing. All testers weren't familiar with this kitchen. For this reason all necessary tools and ingredients were placed around the table (see figure 5.1 below).



Figure 5.1: Kitchen table setup

For the tools a strainer, a measuring cup, a small pot and a pan were needed. Additionally, three bowls, one big knife, one small knife, one turner, one peeler, one scissor, two teaspoons, two tablespoons, two forks and two knives were provided. The sushi burger recipe requires following ingredients for two burgers, see table 5.2.



For the rice	For the filling
230 g sushi rice	1 nori paper
400 ml water	1 tsp wasabi sauce
60 ml rice vinegar	175 g salmon
1 tsp salt	1 small avocado
2 tsp sugar	$\frac{1}{4}$ cucumber
4 tbsp soy sauce	
2 tbsp sesame	

Table 5.2: Sushi burger recipe

During the whole cooking session the user wears the HoloLens. The character was placed around 1m away from their working area, a little above eye height, in front of the cutting board. To wash the rice the tester had to walk eight steps to the right around the working area to a small sink.

A video was recorded throughout the whole test through a DSLR in front of the user. It only recorded the tester's reaction and action. The front view of the camera is seen in figure 5.2.



Figure 5.2: Video recording view of a participant

The DSLR couldn't record what the tester was seeing. This would need a second HoloLens device attached to the camera. Instead, an application called "Microsoft HoloLens<sup>4</sup>" on the windows laptop was connected to the HoloLens and recorded the user's view with the character. The downside of this application is that it records only for a total length of five minutes, then it needs to be triggered manually again. But the application continuously showed that it was recording, even after five minutes. Only a small white light on the left side of the HoloLens indicated the real recording.

<sup>4</sup><https://www.microsoft.com/en-us/store/p/microsoft-hololens/9nblggh4qwnx>

This was not always clear to see. For this reason there were gaps in the video HoloLens recording. The scanning of ingredients was also not possible to record because the application accessed the camera and stopped the video recording.

Before every test, the participant's name was entered into the system so that the character can use the tester's name while interacting. Since the users already knew what will be cooked, they were told to tell the character that they don't know. To ensure that each tester had the same conditions, all testers received the same information as well as instructions on how the user can interact with the character. The tester needed to gaze at the character and listen if a short beeping sound is played. Then the character is listening to what the tester is saying. The air tap gesture was explained and a short instruction on how to use the induction hotplate. Since it was also a qualitative interview, the testers were informed that they could share their thoughts and problems about the application and HoloLens while they are cooking.

None of the testers were informed on what they can say to the character. Only the trigger word to start the application was told, which is "Hello" or "Hi." And all of them were inexperienced in using this application.

Once the application started it guided the tester through the cooking process from giving inspiration what to cook to the actual cooking of the given recipe, the sushi burger. Testers interacted with the application through voice recognition and an air tap gesture. The air tap gesture is used for switching between three different recipe suggestions.

The application can be divided into the following sections:

### **Welcome Message**

The character greets the tester with his name and has some basic level of small talk, directing the user towards what to cook. This interaction between character and tester tries to mediate a personal relationship.

### **Cooking Inspiration**

Since the tester pretends not to know what to cook, the character asks some general questions, like how much time, how difficult the recipe should be, or for how many people the tester wants to cook. As the last step the character requests the tester to put some vegetables on the kitchen table so that it can be scanned. Based on these ingredients the application will give a choice of three different kinds of recipes: a guacamole dip, a salad and a sushi burger. At least two vegetables need to be recognized in order to show the menu. For this test a cucumber, an avocado or a tomato can be recognized.

### **Cooking Guide**

Once the user selects the sushi burger, the character will inform what tools and ingredients are needed. After the tester confirms that everything is prepared, the character will guide the tester through each cooking step. It will show a video, an ingredients list and a step-by-step guide for the current process. The tester can continue each step by gazing at the character and telling it to continue. At the same time, there are buttons below the video to jump to a cooking section. These four buttons give the opportunity to start the section from the beginning when selected. See

Chapter 4 for more information. This is in case something goes wrong or the tester wants to repeat something.

### **Goodbye Message**

To give the character a human touch it ends the cooking session by saying that the sushi burger looks really yummy and the character wants to eat it. After that the character tells the tester that the cooking was tiring and it will take a nap. Here the character tries to connect to the tester again on a human level. After that the application will terminate itself.

To minimize the time for the cooking session, the rice cooking was modified. The rice was still cooked by the user but instead of waiting until the rice was done the cooked rice from the previous session was given once the tester had to flavour the rice (the rice for the first session was cooked by the researcher). Otherwise the tester will have to wait for the rice to cool down.

During the cooking sessions the researcher was standing next to the kitchen table, trying to avoid interference. Once technical or interactional problems occurred the test was interrupted by the researcher. When testers chose to share their thoughts, questions were asked to narrow their thoughts further down. During the cooking of the salmon, which took around 10 minutes, the testers were asked questions regarding functionality of the object recognition and the character (Appendix B.1.1). A few interview questions were asked once the tester finished the sushi burger (Appendix B.1.2). They then had the option to make a second burger and take it home.

Answering the questionnaire (see Appendix B.2) was conducted on a table next to the kitchen setup. While testers were answering the questions, the kitchen was cleaned and prepared for the next test session.

The timer was stopped during three stages: introduction to the character, start of the rice cooking, and during the last step of the sushi burger.

## **5.5 Data Analysis Plan**

During each cooking session first notable observations and think aloud thoughts regarding application and HoloLens have been documented.

After all tests were completed the videos were transcribed. Therefore all videos were watched and notable observations in terms of behaviour, emotions, problems and what they were saying regarding application and device were gathered and listed in an Excel file.

The data analysis method was derived from Mayring, a classical approach to summarize the findings and form categories (Mayring (n.d.)). The elaboration occurred inductive, meaning that the categories were structured while going through the transcript and not before. Several categories have been created and split into more details. The results from the quantitative and qualitative evaluation were added into the respective category. For each category the number of occurrences was counted where it was possible.

The quantitative data from the questionnaire and from the AttrakDiff tool were evaluated separately because the quantitative data of the questionnaire consists of background information about the participant, whereas the AttrakDiff tool was used to assess the overall application. AttrakDiff already includes an analysis of the evaluation.

## 5.6 Legitimation

The prototype developed for this thesis has only basic functionality. Because the application is voice controlled, participants were able to say what they wanted to and it's difficult to predict which words they will choose or what they will say. As many words, sentences and combinations as possible were trained to the LUIS service in order to assist but there was always a chance that the application wouldn't understand the user.

Another limitation was there were instances that the character wouldn't continue for some reason. The researcher had to step in and try to solve the problem or restart the application. In those cases, the researcher went through the application step by step again to the point where the problem occurred and let the participant continue from there. Despite this limitation, the final data would not be affected by this since the research questions were about overall insights and features.

The conducted sessions were all done in a kitchen setup at university. This setup is not the same as someone's home, yet it was accurate enough to answer the research questions. Besides, every participant was unfamiliar with the university kitchen and had the same circumstances as the others.

Due to the issue of lack of time, only one week of cooking sessions were conducted. This limited the number of participants for this study. For a detailed quantitative study more participants would be needed to make a more reliable statement.

## 5.7 Conclusion

The chosen mixed method design consisting of quantitative and qualitative methods is suitable to get insights for the planned cooking session to answer the given research questions. Only the small sample size of participants limits the evidence of the quantitative study to make a more confident statement.

All information gained from the conducted cooking sessions will be described in the next chapter which reveals the findings regarding the HoloLens device and the application.

---

# Results

Previously it was stated how information will be obtained from the cooking session. This chapter will describe the outcomes which were found during the session. This chapter will also give an overview of the participants' background data, as well as share findings regarding the HoloLens device and application.

## 6.1 Participation Profiles

The questionnaire filled up after the cooking session revealed supplementary information about the participants. There were 14 in total: eight female and six male students. All of them have never cooked a sushi burger before the cooking session. On a scale of zero to ten, they rated their cooking skills on average as 6.4, with the lowest number at 4 and the highest number at 8 (SD = 1.22).

In terms of cooking frequency per week, two participants (14.3%) cook twice a week on average; 28.6% cook three days; 28.6% cook four days; and 28.6% cook five days a week. When asked for the reasons why participants don't cook more often, several reasons were named. 40% said it's because they already eat at university. 50% don't have time to cook, and 20% don't know what to cook (see figure 6.1 below).

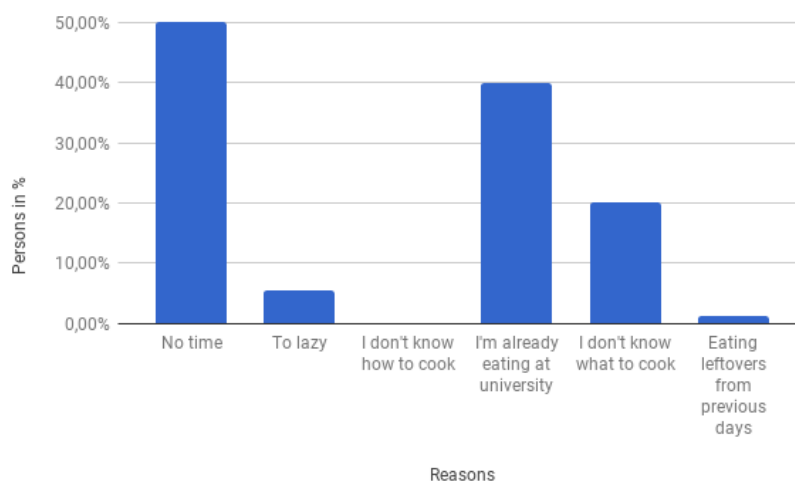


Figure 6.1: Reasons for not cooking more often

If the person does cook at home they don't spend much time doing it. Eleven students indicated that in general, they cook between thirty minutes to one hour. Only one person takes the time to spend up to two hours cooking. On the other hand, two persons say they spend less than thirty minutes for cooking.

Differentiating between cooking something new often and not often, eight persons answered with not often. The reasons for that were mostly "don't know what to cook," "it takes time," and "laziness."

The questionnaire also inquired how the participants learned to cook. They were asked to select a maximum of two answers. Yet, three participants didn't follow the instructions and picked three options. Nevertheless, 11 people learned cooking through parents and/or through websites (illustrated in diagram 6.2).

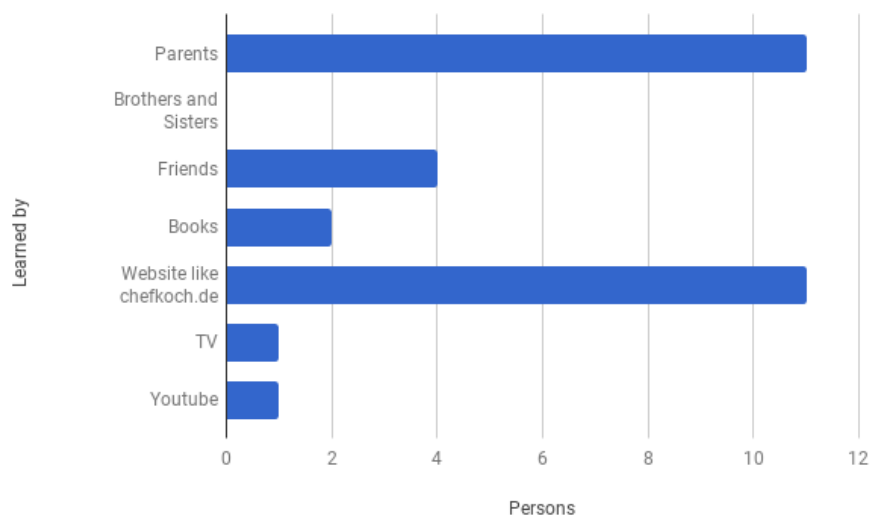


Figure 6.2: Where participants learned to cook

During cooking, the time from the start of the interaction with the character until the burger was finished with the last step was taken. The time ranged between 35 minutes and one hour and five minutes (Mean = 49.07 minutes, SD = 8.5). However, there were several instances where problems occurred and the researcher had to intervene, which created bias in the results.

## 6.2 HoloLens

The first research question asked which restrictions exist in using a hands-free device in the kitchen context. To answer this question, findings through observation, think aloud methodology, and a structured questionnaire were obtained and correlated. Four categories were identified: hands-free, optics, interaction and overall.

### 6.2.1 Hands-free

An interesting fact was revealed through observation regarding hygiene. Overall, eight participants washed their hands, among them seven females. On the other hand, six testers (five males and one

female) cleaned their hands with kitchen paper. The participants were confessing in the questionnaire that they realized that they washed their hands less than when using a tablet/smartphone. Undoubtedly, participants liked that they had their hands free. Eight persons mentioned that fact in particular in the questionnaire.

### 6.2.2 Optics

Problem	Participant														Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Looking below the glasses				x		x			x			x	x		5
Field of view limited	x			x			x					x	x		5
Light reflexion in the glass				x			x							x	3

Table 6.1: Optic problems overview

Another interesting aspect was commented on by five participants and that was the fact that they were looking below the glasses during the cooking process (see table 6.1). During observation it was realized by a few more people. Because the recording does not necessarily represent what the user sees since the sensor is situated above the eyes, it is not immediately obvious that people were looking below. The following response was made by Participant 9: "I'm used to having no glasses since I'm generally not wearing glasses. It's something unfamiliar." Participant 12 said, "The field of view should be further down. I'm somehow afraid that the device is falling off my head that's why I'd rather look below." He shared that he had to move his head completely vertical in order to look down. Tester 13 pointed out: "Through the glasses I can't recognize colours that well." When asked for the reason, he said it's because the HoloLens was darkening the view. In some other cases, the video was overlaying what was being looked at on the kitchen table.

One criticism mentioned by five people was the limited FOV. Two persons observed that they had to step back or to the side during cooking to see the character.

Lastly, light reflexions on the glass was mentioned by three participants even though, everyone had the same lightning conditions from the ceiling. There were windows further away, but those were also blocked by some closets. It's possible that the sun had some minimal influence on this. Participant 7 said that while he was measuring rice it was "a little bit irritating that it is reflecting depending on... . while I'm standing here I can see that the light is breaking... So to say green, red lines." He later noticed that "glasses have a light grey tinting. And then finding details, like opening of the rice vinegar bottle... That is difficult." In addition, Participant 14 observed this while putting salt on the salmon in the pan: "This is a bit difficult because you don't see what you're doing. Currently I have a lot of colourful light reflexions in the view." It also got mentioned in the questionnaire that followed that better see-through glasses were wished for.

### 6.2.3 Comfort

Through observation it was noticeable that ten out of 14 participants (seven females and three males) had to adjust the HoloLens throughout the cooking sessions several times ( $SD = 3.27$ ) (see table 6.2). The device was slipping off of their nose and had to be put back further up repeatedly. "The glasses are apparently for big noses," stated one participant.

Problem	Participant														Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Adjusting HoloLens (observation)	3	7			11	2		1	7	1	3		2	3	40
Device is heavy (interview)	x	x	x	x	x	x	x	x		x	x		x	x	12
Hairs are in the way (observation)		1			1			1							3
Device is warm (interview)							x								1

Table 6.2: Comfort issues overview

Likewise, 12 participants (eight females and four males) mentioned that the device is heavy. An example of a response included: "It's getting heavy over time and is pressing against the nose. For longer preparations that will be uncomfortable." Another person mentioned that she always had the feeling that the device will fall off because it's heavy.

Three female participants also had a problem because their hair fell into the field of view and they had to put it aside during cooking. However, this only happened once per cooking session. To increase the comfort of using the HoloLens, six people suggested making it lighter. Five persons said the general comfort while using the device must be better.

#### 6.2.4 Interaction

A menu selection with three buttons were shown once the vegetable was recognized. In order to start cooking the sushi burger, participants had to select buttons using the air tap gesture. It was briefly demonstrated before cooking but four females and four males faced problems in getting the button selected. Among them were two who had tried the HoloLens before. The third person who had previous experience with the device, didn't face problems. Either the gesture was executed incorrectly, too slowly, successively, or the view of the button wasn't focused. One comment said, "I can see the border around the sushi burger button but I don't know if my fingers are getting recognized." One participant bypassed the gesture by talking to the bot and saying, "I would like to cook the sushi burger." Another tester made first, a side gesture with his hand once he saw the menu, and then the air tap gesture.

#### 6.2.5 Overall Remarks

Among some general feedback from participants regarding the HoloLens which participants gave was that "It takes time to get used to the interaction with the HoloLens. Still a bit unusual." This was mentioned by two persons. A tester said: "For new recipes which you have never made before it might be really funny. But I'm not sure if the device will be disturbing because of its size." On the other hand, one tester said, "I think if you decide to wear the HoloLens while cooking, then you probably already have an idea what you will cook, I would imagine." Another participant remarked that it's a good option compared to a smartphone or tablet.

When asked in the questionnaire about what would prevent the participants from using the HoloLens, four named the high price. Two persons worried about getting headache or that the glasses will get misty. Some mentioned that they would not use it if it will take longer to cook



with the HoloLens than without, or when other people are in the room. Someone likes to listen to the radio while cooking, and she's wondering if the radio could accidentally trigger the next step.

## 6.3 Application

While the previous section demonstrated findings about the HoloLens itself, this section will now focus on the application. For this scope of the thesis an application was developed to guide a user from giving inspiration on what to cook to a step-by-step instruction for cooking a sushi burger. The developed application shows a possible solution on how to assist during cooking in order to answer in what way a hands-free device, coupled with media support, can assist while following new recipes. All information gathered from observation, think aloud methodology, semi structured interview and the structured questionnaire will be now presented in a combined manner, whereas, the results of the AttrakDiff tool which examines the overall application are listed separately. The results of the application were divided into four different aspects: Character, video and text, scanning feature and additional features.

### 6.3.1 Character

An evaluation of the conducted cooking session exposed some key findings regarding the character in terms of appearance, personal relationship, interaction and sound.

#### Appearance

Throughout the cooking session participants interacted with a flying robot head. When asked if they were in favour of having a character or if having just a voice was preferred, eight people (57.1%) liked having a character (see figure 6.3).

Four participants would be OK with just having to listen to a voice, and the rest were not sure. Someone liked the idea that the character specifically needed to be looked at in order to move. Another person mentioned, "Only voice would be creepy. The character makes it more personal and it's good to know that someone is there." Someone else mentioned that "I need to talk to him to continue... If I would be at the fridge and want to continue to see what's next. I don't want to turn around to speak with someone first."

In addition, they got asked what kind of character they would prefer: a human character or something fictional. A majority of nine people tend to prefer having something fictional. They liked the flying robot head, but they also could imagine a comic figure or anything funny. A human was preferred by only one person. In general, the participants didn't like the idea of having someone they know as a human character. They would rather the persons were there for real. If it would be a human form, it needed to be really realistic.

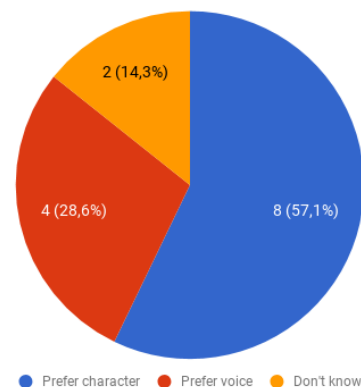


Figure 6.3: Character or voice preference

When asked if they liked the idea of having a character in the form of a known chef like Jamie Oliver, five testers said yes. However, someone mentioned that if that were the case, then the application should only suggest recipes by that particular chef. Another person liked the idea because when a chef is the one giving instructions, then the end result must be something good.

Another aspect mentioned by three users is that they like that the character stays in one place. For three people it occurred that the character moved to a higher position during the cooking session. This mostly happened when they were coming back from the sink, which is at a darker spot. The HoloLens somehow lost the connection or was scanning the area at the small sink. When the user went back to the kitchen table, the character was at a different position.

### Personal Relationship and Emotions

Key findings	Feedback
Character greets user by name	5 smiling
Character misunderstands user	6 laughing
Character says it looks yummy	1 annoyed
Character says good job	7 smiling
Character says its taking a nap	9 grin

Table 6.3: Triggered emotions findings

In the beginning and at the end of the application the character tries to build a personal relationship with the user. When starting the application, the character greets the user with their name. Five out of 14 users responded with a smile as shown in table 6.3 above. Even more participants, six in total, laughed when the character misunderstood them. At one point in the application, the character would repeat what the testers said and it was sometimes different from what they were saying.

When the sushi burger was finished, the character would mention that the sushi burger looked yummy. One tester found this ridiculous since he knew the character couldn't actually see the burger. To further compliment the user, the character would tell the tester that he did a good job, followed by the tester's name. Seven users were smiling as they heard that. Afterwards, the character announced that he will take a nap since cooking was tiring. As a result, nine participants had a grin on their face.

Half of the participants stated in the semi-structured interview that they liked the human touch. Someone stated that it can be helpful for someone who cooks alone. Another person remarked that during the cooking the personal relationship didn't matter, but that he felt somewhat proud when the character said that he did a good job. However, the character still has room for improvement to establish a connection.

### Interaction

All participants had problems interacting with the character. In order for the character to listen, testers had to gaze at it and wait to hear a short beeping sound. The sound was triggered when they looked away and then looked at the character again. The beeping sound indicates that the character is listening. Most of the time, users replied directly after the character was finished

speaking without hearing the beeping sound. Some users needed additional instructions while others figured it out after some time even though the information on how to talk to the character was given before the cooking session started. Three participants also tried with slower and better pronunciation when they repeated what they were saying.

Another aspect was interrupting the character while it was speaking, which was observed with two persons. They wanted to move to the next step before the character was finished talking.

Always having to look at the character and say okay was annoying for one participant. Another participant mentioned that he would rather have a menu where he can trigger the next step than always talking to the character. He assumes that this would feel more convenient and faster.

Lastly, two testers tried to touch the character. They were reaching out with their hand and wanted to grab it. It occurred when one person was in the process of starting the application, while the other instance was when a person was trying to move the character when it was in the way.

## Sound

The character made a constant whirring sound. The sound was at a minimum volume. Despite this, six participants were pointing out it's annoying. Two of them felt even pressured by this noise. The short beep sound which the character makes while looking at it is mentioned as annoying by one person.

### 6.3.2 Video and Text

The guidance through the recipe was through text and video. There was one common problem which everyone faced. Once video and text were displayed it occurred that the character was overlapped by text or video. This happened when the user moved their head around because video and text were following the movement (figure 6.4).

During implementation and testing it didn't seem like a big problem. After the first few tests it showed that this was annoying for all of them. But because all testers should have the same condition, nothing was adjusted in the application.

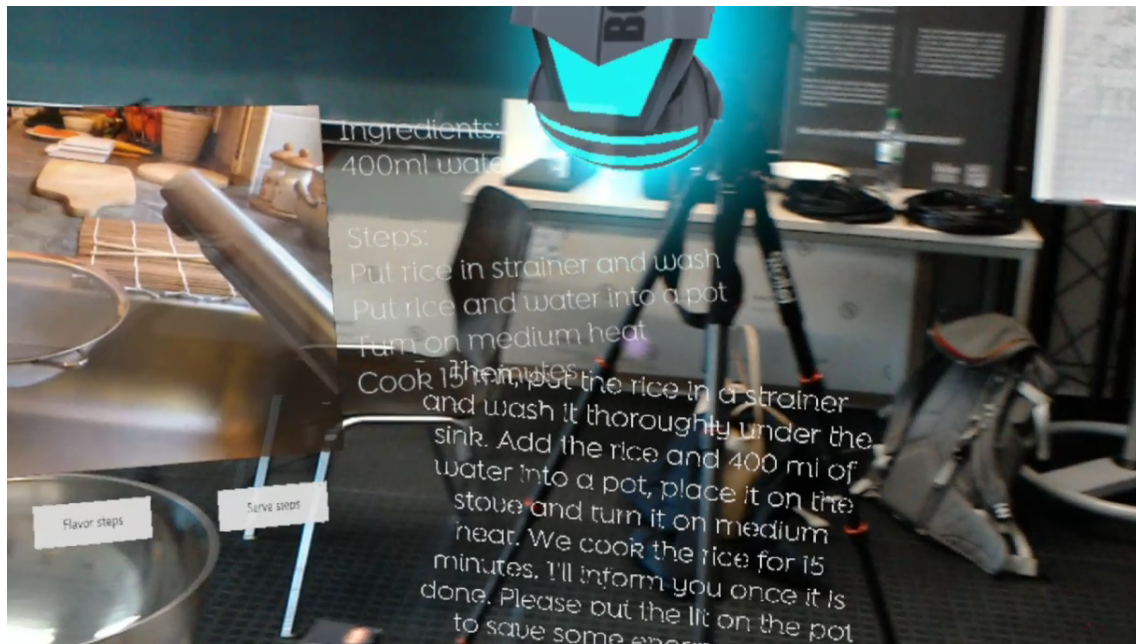


Figure 6.4: Text and character overlapping

An additional request was made by two students. They found it disturbing to see the text and video all the time, even when walking around. It would be nice if both could be deactivated and made visible again when needed.

Observing how accurate participants followed the instructions, it was revealed that only one person followed instructions as it was shown in the video and in the text. All others made at least minor adjustments like cutting things differently, not accurately measuring oil or soya sauce, adding too much of soya sauce, or changing the order of steps.

## Video

Comments	Count
Video is in the line of sight	5
Repeating of video gets annoying	2
Waiting for the video to start over	2

Table 6.4: Video issues findings

All participants except one mentioned in the questionnaire that the video was helpful. Observation showed that all students likewise had more than one closer look at the videos during the cooking process.

A few testers mentioned that the video was in their line of sight while doing something or it was covering what they were looking for (see table 6.4). The repetition of the video sequence could get annoying, especially while two testers were waiting for the fish to be cooked, which took a while. They wanted to have the possibility of pausing the video. Some others had to wait for the video to start over again because they missed the steps in the beginning of the video. One participant was

annoyed because he was doing the last step of the sequence already and had to wait every time for the video to finish.

### Text

Comments	Count
Ingredients and steps are helpful	11
Text is flickering	5
Character text is not needed	4
More details for steps needed	2

Table 6.5: Text issues findings

The text which shows ingredients and overview for the current step was found useful by eleven persons (see table 6.5). Among them, eight students mentioned in particular the ingredients list, whereas six persons liked the step overview.

Some people had problems reading the text because it was kind of flickering. In a specific angle the text appears to be shaking.

In contrast, a few students pointed out that the text is redundant for indicating what the character or the user is saying. Someone mentioned that because the character is speaking so slow and clear, it is no longer needed. Another one said it's like subtitles in a movies. If the character speaks his mother tongue he probably wouldn't need the subtitles at all.

### 6.3.3 Scanning Feature

The character's instructions for scanning the vegetable led to confusion. Only five participants were putting the avocado, cucumber, fish or nori paper on or next to the cutting board for scanning, which was the correct procedure. Five testers just looked at the vegetables as they were arranged on the kitchen table. In contrast, the four users held the cucumber or the avocado in the air and looked at it while scanning.

Asking testers regarding the feature of recognizing vegetables with the application, half of them found it useful. One person was sceptical about it, saying that "you normally don't put everything on the table what you need for the recipe." Four people pointed out that it would be nice to get suggestions for what can be cooked from leftover ingredients.

One participant remarked that "it would be nice if the application could scan the kitchen table and automatically let you know what is still missing for this recipe. Instead of comparing every ingredient with what is displayed in the video or text." Another person would like for meat to be scanned while it is in the pan, informing the user when it needs to be turned or is done.

### 6.3.4 Additional Features

Some additional features were suggested through the interview and questionnaire.

- Four testers requested an integrated timer. They want to get reminded when something is done with timer counting backwards.

- A timeline for the whole cooking process was requested by two participants. To them, it would be nice to see how much progress is being made and how much is left. They also want the application to show which steps can be done simultaneously.
- One tester wanted to see the end result projected on the real plate before the start of cooking.
- One participant pointed out that a back button for cooking steps would be helpful.

### 6.3.5 Results About the Overall Application

Quantitative data of the overall application was measured by the AttrakDiff tool. This evaluation of the application shows that the pragmatic quality (PQ) was rated with an average value overall (see figure 6.5). There is still room for improvement to achieve the tester's goal in an easier way.

At the same time the hedonic quality (HQ) does not clearly apply to a specific character classification because the confidence rectangle is situated between two characteristics. The user is stimulated by the application but the hedonic value is only a bit above average. The confidence rectangle, which is light blue, is small. This is an advantage because it means that the results are less coincidental and more reliable. The users didn't have much disagreements in their rating of the application.

To deeper understand the hedonic quality, it can be distinguished into the aspects of stimulation (HQ-S) and identity (HQ-I). The stimulation aspect received above-average, whereas the identity aspect was rated average (see figure 6.6). The identity factor needs to improve to bind the user more strongly to the application.

As mentioned earlier, the PQ aspect is just average. The overall impression (ATT) of the product is above-average and is very attractive.

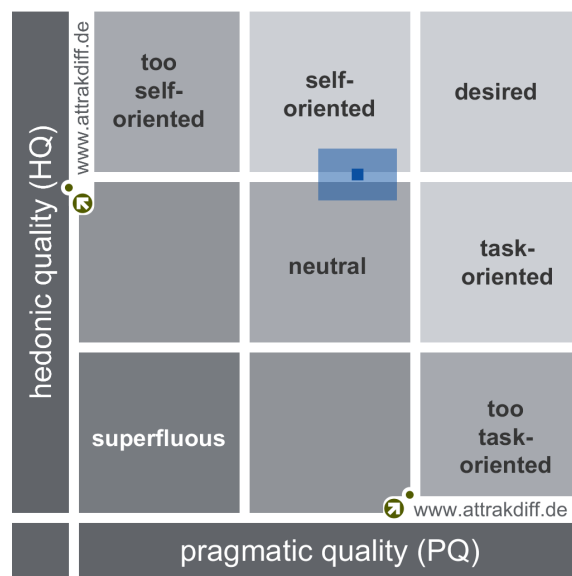


Figure 6.5: AttrakDiff - Portfolio diagramm

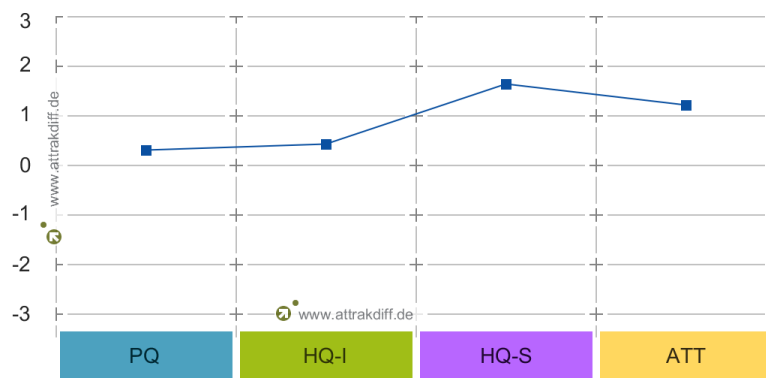


Figure 6.6: AttrakDiff - Diagram of average values

To individually analyse each dimension of the AttrakDiff, an overview with word-pairs is provided, indicating the tester's perception of the application (see figure 6.7).

In Pragmatic Quality (PQ), the application was perceived more technical than human. The highest value in this dimension was received for being a practicable application. Even though the application is practicable, the structure of it tends to lead to confusion since the rating is below average.

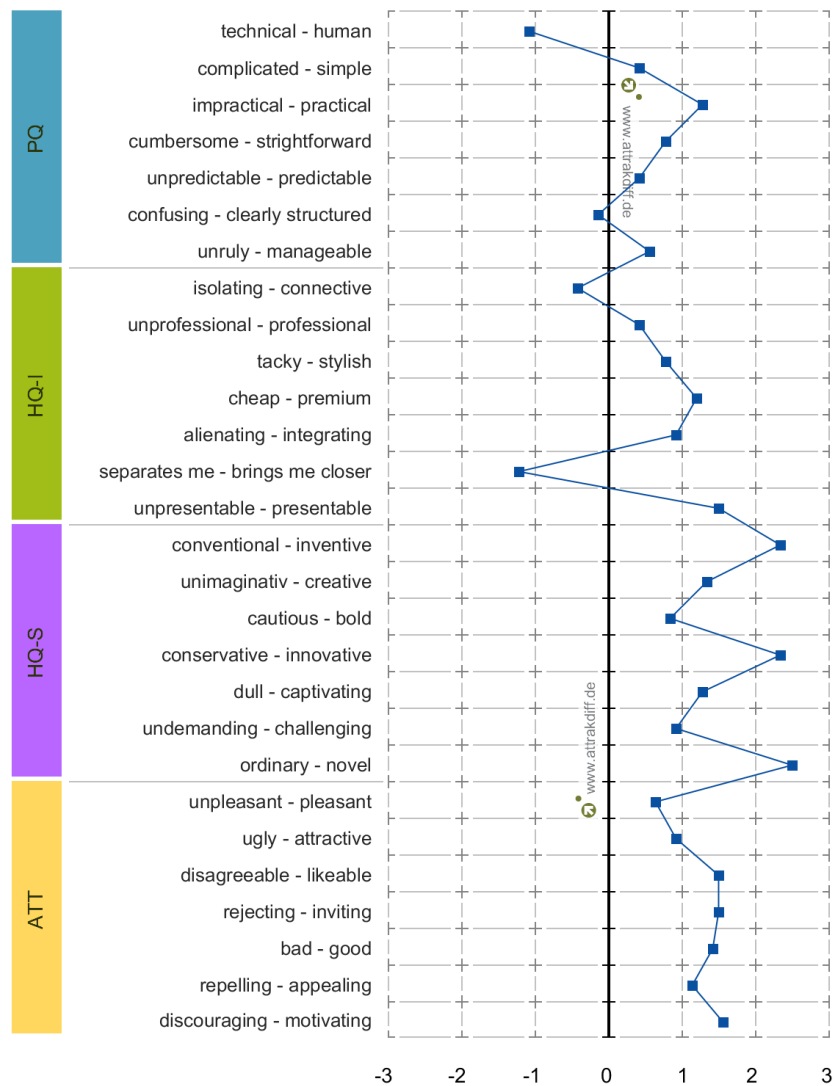


Figure 6.7: AttrakDiff - Description of word-pairs

In Hedonic Quality-Identity (HQ-I) the overall lowest score for this application was achieved for the word-pair "separates me" and "brings me closer." While using the application the user had the feeling that it isolated them. This is also validated through a low score for being connective. Nevertheless, it was being rated as a presentable application.

In terms of Hedonic Quality-Stimulation (HQ-S) the application received an overall above average rating. Three pairings received similar high values: inventive, innovative and novel.

For Attraction (ATT) all values are as well above average.

## 6.4 Conclusion

The applied research methods revealed interesting perspectives regarding the device and application usage. First, the hands-free device led users to be less hygienic. There also seemed to be problems with the optics under certain conditions. This was observed and mentioned in the questionnaire and interview. Whereas for the application users preferred to have some kind of a character to interact with. The character combined with emotions is perceived very well, as proven by the reaction during observation. An important help is seeing a video of the current step, while having a list of the ingredients and summarized steps as visual aid. This will lead to a deeper discussion of the findings in the next chapter.



# Discussion

The previous chapter listed characteristics of participants, as well as findings regarding device and application from the conducted cooking session. Now, how these findings could be interpreted and explained will be discussed with the aim of answering the research questions, the findings are divided into two sections: HoloLens and application.

## 7.1 Evaluation Of the HoloLens

With regard to the research question which restrictions currently exist in using a HoloLens device in the kitchen context, the study contributes insights which will be discussed next.

The findings suggest that the fact that the HoloLens enables the user to interact hands-free was perceived by most of the testers as an advantage. Compared to other ways of following a recipe like using a book, table/smartphone or a print out, it gives the advantage of moving to the next step without touching anything and not switching between doing something and reading. People were reporting that they have never had gone for so long without "dirty hands." People are forced to wash their hands when using a book, tablet/smartphone or printout if they don't want to make these things dirty. As mentioned in the results, six testers were cleaning their hands with kitchen paper. The question arises that if they use a non hands-free device, will they wash their hands instead? Therefore further studies need to be conducted as well as using recipes where hands get dirtier.

Occasionally, testers would confess that they were looking below the actual glasses because the glasses are interfering with the view. The glasses of the HoloLens seem to darken the view, according to testers. Under bright conditions this is not such a substantial problem. However, it can cause serious problems under poor lightning or for people with bad eyesight. If a user looks through the glasses in bad lighting while cutting something then it can cause injuries because the user can't see that clearly. The same can occur when the light reflects on the glasses, although these reflections can occur randomly, influenced by the surrounding light. These problems need to be deeper analysed under bad conditions or next-to-bright sunlight for the reflections. Furthermore, the recipe didn't involve a lot of hot cooking. Glasses might get misty like normal glasses. Whether the electronics can handle humid conditions for a long period of time can't be said.

The FOV is still not enough for some users, similar to the findings from Blattgerste et al. (2017) and Werrlich et al. (2017). If users want to see more of the character or video they need to take a step back or to the side. Otherwise the problem is that the displayed content gets cut off.

Depending on the kitchen setup, this can be tricky when there is little space to move around. Some application content could be adjusted to make things smaller so that it fits into the FOV. But text still needs to be readable, so there the settings are limited. Although, this may not lead to the best user experience, the application is still usable.

Comfort plays an important role as well. The majority of testers had to adjust the HoloLens on their head a few times during the session. It seems that the device was slipping on their nose. Likely this is caused by the weight of the device. This seems to be an annoying factor and needs some improvement. Wearing the HoloLens for longer cooking sessions can be dissatisfying. Some users also had the fear that the device will fall off while they were in the process of doing something and that if both hands were dirty and they needed to touch the device, the device will get dirty as well. Therefore, the electronics and sensors need to be resistant. Nobody wants to have a device which can't be used after a while anymore.

For some people, their hair falls into their view, especially female participants. This happened once per session. It is an annoying factor but since it occurred only a few times it is less annoying than the need for adjusting the device on the head.

Another aspect is the device had a tendency to get warm. Interestingly, this only got noticed by one person. More cooking sessions are needed to verify this problem. Furthermore, it would be interesting to know how the device will perform in a hot kitchen and hot temperature, and if the device will easily adjust to the temperature of the surroundings or get to a level that it is no longer acceptable to wear.

Due to the lack of experience from participants with the HoloLens device, many found that the interaction with menu buttons was not easy to accomplish. This feedback was also given from tests conducted by Werrlich et al. (2017). It took some time until participants figured it out. But once they become used to the gesture this shouldn't be a bigger problem. Despite this, it would be an advantage to select things not only by gestures but with voice as well. In the kitchen environment it could always happen that both hands are needed for other things.

The additional remarks indicated that people wouldn't always use such a device. It needs to get lighter, smaller and shouldn't cause dizziness, even though during cooking nobody had an issue with this feeling. People would tend to use it for recipes which they don't know and probably when there is already an idea about what to cook. It seems like the device wouldn't be used for daily cooking, only when someone is unfamiliar with the steps of a recipe. In addition, participants would use the device only when they are alone. The device would be too distracting if someone else is around during cooking. They want to speak to the character and the other person in the room but since the other person would not see the character, it would be weird, too.

## 7.2 Evaluation Of the Application

The question of how a hands-free device can assist while following new recipes also needs to be answered.

Participants prefer to have a character slightly more than only hearing a voice. Since no previous experience with a voice-only application exists with participants this question can be only answered speculatively. An assumption for why a character is preferable is that otherwise the application is listening to every word spoken, even if only a keyword needs to be analysed. That in

mind will raise privacy concerns for users. It also might happen that the application is responding even when it shouldn't, since the application is always listening. Another possibility is that having a virtual object to interact with feels more natural for people since it's a normal way of interacting with someone. It might also be more difficult to build a personal relationship between a mere voice and the user. Apart from that, how the character behaves and its appearance can be further improved.

The preference for having a familiar person or human character was low. It really needs to be realistic to be accepted, and users would rather have the real person in the kitchen. It will be already sufficient if the robotic head laughs or shows more visual emotions. A comic figure might have also been better accepted than the robot head. The idea of a totally strange human character is disliked because it would give the feeling that a stranger is observing the user. On the other hand, people liked the idea of having a real chef, for example Jamie Oliver, giving them instructions. But likewise the user should only be able to cook recipes from this chef. If a chef is telling them what to do then it must be true. This might raise the confidence level of users in cooking a new recipe.

Another discussion involves the position of the character. For this thesis it was decided that the character stays at one position. Unfortunately it sometimes happened that it moved to a different position due to some technical problems. However, that the character stays at one place generally got accepted. But it does have the disadvantage that when moving around, the user needed to turn around in order to look and talk to the character. Since movements in a kitchen are minimal, this constraint should be manageable. It is only time consuming and might get annoying.

Building a personal relationship with the character won't lead to a worse experience than without one. Only if the character knows too much about the user will it probably get disliked. Praising someone's cooking skills is good but it needs to feel realistic. If the user knows the character can't see the final result but says that it looks good, that's will probably not be liked.

The evaluation of AttrakDiff is showing that the application is perceived as technical. This result could be driven from the robotic voice and the appearance of the character, since it's a flying robot head. But it can also be improved by giving the character more emotional feedback and more knowledge about the user. If the character knows what food preferences the user has or gives the user diverse food choices based on the food the user had during the week as examples, it would build a deeper relationship. Another obstacle is that the application is rated as being very separating and isolating. If more than one person is cooking, everyone should have a device and an application. If only one person is using a device, the other person can neither see nor hear what the application is instructing. When using more than one device the application needs to be adjusted to support multiple users, and the application itself needs to be modified for that. A different analysis needs to be made for how the device will assist when other people with a HoloLens are in the room while cooking. Is there a way to involve more people in the cooking process?

The interaction with the character was a little tricky for some people. They didn't always wait until the beeping sound was played or wanted to interrupt the character during speaking. A solution would be to give proper instructions, in the form of a tutorial, within the application when used for the first time. Since some testers wanted to touch the character, it would be possible to add some interaction for this case. Another problem which arises quite often is the accidental triggering of the character. Some users gave a quick look at the character and the listening got triggered. The

accidental triggering of the character could be improved that it only gets triggered when the user gazes at least two seconds at the character.

Playing a constant sound from the character is not necessary. People will know that someone is there to talk to. Yet, it's necessary to play a beeping sound when the user is gazing at the character. There needs to be an indicator that the character is listening. Without that feedback it would be only possible to know through something visual that the character is listening, but that is easier to overlook.

Video and text are seen overall as helpful. The only problem which needs to be solved is the overlaying of text and video with the character. That might also explain why users gave the application in AttrakDiff a slightly below average value for being confusing. First of all, the text display of what is spoken by the character seems not to be needed. This can be made optional like a subtitle in movies. This change can already make a significant improvement. Second, it would be an advantage if the character will recognize that a text or video is colliding with it. The character could move its position a little and return to the original position when the collision is over.

In addition, it can get disturbing to have the video and text visible the entire time during steps that take a long time to finish. There should be an option to make them disappear and appear again when needed. This could be achieved through a specific gesture or special voice command. Some extra features for the video would be nice like fast forwarding or stopping but are not as necessary. Also, the steps which are displayed in a video sequence should be reduced to a minimum.

The scanning of ingredients feature was accepted by half of the participants. In any case, the user needs proper instructions to use it properly. Since users had issues during the cooking session, the feature did not get perceived as really working. Also in some cases it recognized the ingredients immediately, which was too fast for participants. The threshold for the object recognition was set very low to guarantee a recognition at all. Testers were saying that the recipe suggestion could help for leftover ingredients. Therefore more vegetables needs to be detectable in order to suggest more recipes and be perceived as useful.

A feature which users were missing was the timer function. For the cooking session the researcher acted as a timer. It is enough if the application could show a simple text as a countdown. It would also be helpful in seeing the overall cooking process. Because users are cooking something unfamiliar they can't estimate how long they need for the recipe.

Initially, while recipes are suggested or before the actual cooking takes place, it would be good to display the end result on a plate to show how it is supposed to look like. This leaves the user not totally in the dark as to how the dish should turn out.

# Conclusion and Future Work

The discussion chapter reviewed the findings about the device and application, as well as examined them. With these gained knowledge this chapter will now answer the research questions and inform about possible prospective research which can be built upon this research.

## 8.1 Conclusions

The aim of this thesis was to investigate the usefulness of an HMD device, in combination with an application, while following a unfamiliar recipe. Therefore a prototype application for the HoloLens got developed and tested by 14 participants. The application guides a user with the help of a character through the cooking of a sushi burger by using multimedia content. In a mixed method design study it then got investigated and evaluated.

The following findings in regard to the research questions were found:

### Research Question 1

The first research question aimed to answer which restrictions currently exist in using a HoloLens device in the kitchen context. Undoubtedly, limitations of the HoloLens device are especially the heavy weight and small FOV. However, the weight of the device seems to be more of an issue than the FOV. The view of the device is limited but nevertheless this only restricts how much the user is seeing. By moving the head or changing position, the view can be adjusted to the user's needs, whereas the longer the cooking will take, the heavier the device will feel on the head. This also leads to the problem that some people needed to adjust the device several times during a cooking session.

Another constraint in using the device is the perception of view. Looking through the glasses feels unfamiliar for some or even darkens the view. For this reason people tend to look below the glasses. The problem itself is not only that people don't see what the application would show, but that the HoloLens limits the normal view perception. Issues with light reflexions in the glass can also occur. That can be dangerous in the kitchen, for example when things are being cut, since that suddenly happens under random circumstances and distorts the view.

When a user cooks with the HoloLens and there are other people in the kitchen without a device as well, then these people can't be involved in the cooking process because they don't know what the cooking person is seeing or hearing.

### Research Question 2

The second research question intends to answer in what way a hands-free device, coupled with media support, can assist while following a new recipe. The results of this study showed that an application can be absolutely helpful. Using voice as the central interaction method during cooking is of advantage. Even applying gestures works seamlessly because there is no need to touch anything directly. Video and text for steps is certainly recommended even though the steps are mostly not followed 100% as instructed. Since the steps are always present, the user doesn't lose the context and always knows which step needs to be executed. A character guiding through the steps appears to be an accepted choice. Results also showed that it is possible to build a relationship between user and character to trigger user's emotions, even if the application is perceived as being technical. These positive emotions might encourage the user to cook more often.

To summarize, the HoloLens device itself has some limitations but it can already be used in the kitchen to assist during cooking these days, especially for recipes which don't involve too much hot cooking. Together with an application which shows videos and text for every single step as well as being interactive through voice and gestures, it can serve as a better learning method for unfamiliar recipes, compared to following recipes through the old way, using books, phones or tablets.

## 8.2 Future Work

Since no research has been made in the area of using a HoloLens for cooking at the time of writing, there are still plenty of opportunities for future research. This thesis only focused on guiding through steps of making a sushi burger with a HoloLens.

Before conducting further research with this application, the application will need some adjustments so it is less error prone, especially in the interaction with the character and the guidance through the application. One possible next step would be conducting a research where a comparison is taking place between someone cooking with HoloLens and someone without. This will be useful in distinguishing if it will save time at all. At the same time it could be also investigated how-error prone the different approaches are. Will the HoloLens reduce the number of mistakes being made? Another research should investigate the device when something more complex is cooked that involves steaming, which will mist up the glasses. In short, to investigate if such a device can help under more challenging situations.

Further enhancements for the application can be continuously added such as the following ideas: having a famous chef as a character, giving the character more knowledge about the user, or an overall improvement of the character assistant. The application could make a weekend plan based on preferences or previously cooked dishes.

Once the device gets lighter and less bulky it would be interesting to test this device and application with early-middle stage cognitive impaired people. Dementia is becoming an increasing worldwide problem and no cure is available so far. In an updated version the application could also help blind people with cooking. It could use the object recognition feature of the application and communicate what the HoloLens is seeing solely through audio. The application could tell the user where things are located, warn if something is boiling or just say the current step. Whereas blind people cannot see, others need visualization as help. It needs to be explored if it would help

these people to live more independently so that they could cook on their own. It would be a small step but that would tremendously increase life quality. The ability to do something independently and be useful makes life much more liveable for the individual.

---

# Cooking Flow

A sample flow how the user can interact with the character.

## A.1 Welcome Flow

User	Hi
Character	Hi <i>user</i> , how are you today?
User	I'm fine.
Character	That's always a good start for cooking! So are you hungry?
User	Yes, I am.
Character	I guess that's why you called me. So, how can I help? Do you know what you would like to cook today?
User	No, I have no idea.

## A.2 Cooking Inspiration

Character	Okay. Before we start cooking, I am going to ask you some questions that I know what you want. How much time do you have?
User	Around an hour.
Character	OK got it, thank you. Do you want to cook something easy or more difficult today?
User	Something easy.
Character	Something easy, I see. For how many people do you want to cook today?
User	Two persons.
Character	Do you have some special requests for today?
User	No.
Character	OK. If you place some food on the working space, I can help you with some ideas. Please put some vegetables into the grey area which will appear and hold your view steady. I will suggest a recipe once I found at least two ingredients.
User	<i>The user puts ingredients on the surface and it gets detected. In the menu suggestion, the user selects the sushi burger recipe.</i>
Character	Great choice. This meal sounds yummy to me :) <sup>1</sup> . Let's start in a few seconds.



### A.3 Cooking Guide

Character	One more thing. After each step you need to look at me and say OK or next, in order to reach the next step. Say ok, if you got it.
User	OK.
Character	Okay, let's start. You need 1 strainer, a cutting board, 1 pot and 1 pan.
User	Next.
Character	For the burger, prepare sushi rice, rice vinegar, sesame, salt and sugar. For the filling of the burger you'll need nori sheet, salmon, avocado, cucumber, soy sauce, wasabi sauce, mayonnaise, some salt and pepper.
User	OK.
Character	First, weight 230 gram of sushi rice with the pot. That's 200 gram sugar on the measuring cup.
User	OK.
Character	Then, put the rice in a strainer and wash it thoroughly under the sink. Add the rice and 300 ml of water into a pot, place it on the stove and turn it on medium heat. We cook the rice for 15 minutes. I'll inform you once it is done.
User	OK.
Character	For the filling, let's start with the avocado. Take a cutting board and a knife. Divide the avocado into 2 halves and remove the seed. Cut it into small pieces.
User	OK.
Character	Then take 1/4 of the cucumber and peel it. Also cut it into small pieces.
User	OK.
Character	Take a middle-sized pan for the salmon. Flavour it with salt and pepper on both sides. Put the pan on the stove, add some oil and wait until it's hot. Then roast the salmon gently on medium heat for 4 minutes on each side.
User	OK.
Character	We can now flavour the rice. Add 40 ml of rice vinegar, 1 teaspoon of sugar and 1 teaspoon of salt into the pot with the rice and mix it. After you have done that, let the rice cool down.
User	OK.
Character	Then, take the roasted salmon, put it in a big bowl and chop it into small pieces with a fork. Add 4 tablespoons of soy sauce to give it some extra flavour.
User	OK.
Character	We can now make the sauce for the burger. Put 1 teaspoon of wasabi sauce and 4 tablespoons of mayonnaise in a bowl and mix it together.
User	OK.
Character	Carefully cut the nori sheet into 4 pieces. You can cut as many pieces as you want to have burgers. For 2 persons, 4 pieces should be fine.
User	OK.
Character	Take the small bowls and lay cling foil into it. Take a handful of rice and put it into the bowl to form a bun.

User	OK.
Character	Now we are ready to make the burgers. Carefully turn the bowl upside down and put the rice on the cutting board. Remove the foil.
User	OK.
Character	Now you can put the nori sheet and filling onto it (avocado, cucumber, salmon). Add as much filling as you want to have in your burger. Then add the wasabi sauce. Finally, put a second bun on the top. Now you only have to add some sesame at the top. That's it!
User	OK.

#### A.4 Goodbye Message

Character	You made your first burger. If you want to make more, go for it. I'd really like to eat that, looks really good. Good job, <i>user</i> ! If you don't need my help anymore just tell me that you're done. Otherwise just select a button if you need to repeat some steps.
User	I'm done.
Character	Enjoy your meal. I'll take a nap now, cooking was really tiring. Call me when you need help again with cooking. See ya, <i>user</i> .
User	Goodbye.

# Questions

## B.1 Semi-structured Interview

### B.1.1 During Salmon Cooking

- Did you find the object recognition useful?
- Do you think only hearing a voice would be irritating? Without seeing a character.
- What would be if the character would be a person who you know? Someone you trust.
- And would walk on the kitchen table?

### B.1.2 After the Sushi Burger Is Done

- Did you find the application useful?
- Did you find the interpersonally of the character useful?

## B.2 Questionnaire

### B.2.1 Background Information

- What is your name?
- What is your gender?
- What is your age?
- Have you ever tried the HoloLens before?
- How many days do you normally cook in a week?
- What are the reasons you don't cook more often?
- How much time do you spend for cooking in average?
- How often do you try to cook something new?
- If not often or never what is stopping you?

- What is the main reason you cook at home?
- Where did you learn to cook?
- For how many persons do you cook in general?
- On a scale of 0 to 10. How experienced are you in cooking?
- Have you ever cooked a sushi burger before?

### B.2.2 Questions Regarding Application

- What features of the application were helpful during cooking? Which weren't?
- What features of the application would need to be improved to make it more useful?

### B.2.3 Questions Regarding HoloLens

- What features of the HoloLens were helpful during cooking?
- What features of the HoloLens would need to be improved to make it more useful?
- What challenges did you face the last time you tried out a new recipe (without the HoloLens)?  
How did you read the recipe?
- How do you think the HoloLens could have helped in that situation?
- What reasons/factors would prevent you from using the HoloLens during cooking?
- If you had money would you rather buy smart kitchen devices (connected fridge, table projector, Thermomix,...) or a smart glass for use in the kitchen?

# Bibliography

- Annunziata, M. and Abraham, M. (2017), ‘*Augmented Reality Is Already Improving Worker Performance*’. [online] Available at: <https://hbr.org/2017/03/augmented-reality-is-already-improving-worker-performance> [Accessed 02 Jan 2018].
- Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S. and MacIntyre, B. (2001), ‘*Recent advances in augmented reality*’, *IEEE Computer Graphics and Applications* **21**(6), 34–47.
- Azuma, R. T. (1997), ‘*A Survey of Augmented Reality*’, *Presence: Teleoperators and Virtual Environments* (4), 355–385.
- Bardi, J. (2017), ‘*ARKit, ARCore, and the future of Augmented Reality*’. [online] Available at: <https://www.marxentlabs.com/augmented-reality-future> [Accessed 02 Jan 2018].
- Beau, Zeller, M., Rwinj and Bray, B. (2018), ‘*What is a hologram?*’. [online] Available at: <https://docs.microsoft.com/en-us/windows/mixed-reality/hologram> [Accessed 22 Mar 2018].
- Blattgerste, J., Strenge, B., Renner, P., Pfeiffer, T. and Essig, K. (2017), *Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks*, in Unknown, ed., ‘*Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '17*’, ACM Press, New York, New York, USA, pp. 75–82.
- Bottéro, J. (1985), ‘*The Cuisine of Ancient Mesopotamia*’, *The Biblical Archaeologist* **48**(1), 36–47.
- Bray, B. and Zeller, M. (2018), ‘*What is mixed reality?*’. [online] Available at: <https://docs.microsoft.com/en-us/windows/mixed-reality/mixed-reality> [Accessed 02 Apr 2018].
- Buche, M. (2017), ‘*ARKit and ARCore Prove Augmented Reality Is Here To Stay*’. [online] Available at: <https://gravityjack.com/news/arkit-arcore-augmented-reality/> [Accessed 05 Jan 2018].
- Buykx, L. and Petrie, H. (2012), *Recipe Sub-goals and Graphs: An Evaluation by Cooks*, in M. Sano and I. Ide, eds, ‘*Proceedings of the ACM multimedia 2012 workshop on Multimedia for cooking and eating activities - CEA '12*’, ACM Press, New York, New York, USA, p. 59.

- Concept kitchen 2025* (2015). [online] Available at: <http://conceptkitchen2025.ideo.london/> [Accessed 10 Jan 2018].
- European Patent Office (2018), '*Alex Kipman - HoloLens: mixed reality smart glasses*'. [online] Available at: <http://youtu.be/YvOnZW4nAuQ?t=4m39s> [Accessed 02 May 2018].
- Fink, C. (2018), *Charlie fink's metaverse: An AR enabled guide to VR & AR*, 1st edition edn, Cool Blue Media, Washington DC.
- Google (n.d.), '*Tech specs*'. [online] Available at: <https://support.google.com/glass/answer/3064128?hl=en> [Accessed 09 Jan 2018].
- Google Glass (2015), '*We're graduating from Google[x] labs*'. [online] Available at: <https://plus.google.com/+GoogleGlass/posts/9uiwXY42tvc> [Accessed 16 Jan 2018].
- Gorey, C. (2017), '*ARCore v ARKit: Google and Apple's battle over the AR world has begun*'. [online] Available at: <https://www.siliconrepublic.com/machines/arcore-arkit-ar-google-apple> [Accessed 05 Jan 2018].
- Hassenzahl, M., Burmester, M. and Koller, F. (n.d.), '*Der User Experience auf der Spur Zum Einsatz von www.attrakdiff.de*'.
- HelloFresh (2018), '*Geschäftsbericht für 2017*'. [online] Available at: <http://ir.hellofreshgroup.com/download/companies/hellofresh/Annual%20Reports/DE000A161408-JA-2017-EQ-D-00.pdf> [Accessed 02 May 2018].
- HelloFresh (n.d.), '*Simply Delicious Meals*'. [online] Available at: <https://www.hellofresh.com/how-it-works/> [Accessed 23 Apr 2018].
- Heyder, T. (2016), '*Guided cooking: Consumers are embracing the convenience of "smart" kitchen appliances*'. [online] Available at: <https://blog.gfk.com/2016/04/guided-cooking-consumers-are-embracing-the-convenience-of-smart-kitchen-appliances/> [Accessed 20 Jan 2018].
- Horie, A., Mega, S. and Uehara, K. (2006), *The Interactive Cooking Support System in Mixed Reality Environment*, in '2006 IEEE International conference on multimedia and expo ICME 2006', [IEEE Operations Center], [Piscataway, NJ], pp. 657–660.
- IfD Allensbach (2017), '*Anzahl der Personen in Deutschland, die kochen, nach Regelmäßigkeit von 2013 bis 2017 (in Millionen)*'. [online] Available at: <https://de.statista.com/statistik/daten/studie/171172/umfrage/haeufigkeit-vom-kochen/> [Accessed 21 Jan 2018].
- International Organization for Standardization (Mar 2010), '*Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems*'. [online] Available at: <https://www.iso.org/standard/52075.html>
- Kipmann, A. (2017), '*Happy birthday, HoloLens! A huge thank you to our community*'. [online] Available at: <https://blogs.windows.com/devices/2017/03/30/>

- happy-birthday-hololens-huge-thank-community/#urJFGzctei7VCCLR.97 [Accessed 10 May 2018].
- Kreylos, O. (2015), '*On the road for VR: Microsoft HoloLens at Build 2015, San Francisco*'. [online] Available at: <http://doc-ok.org/?p=1223> [Accessed 06 Jan 2018].
- Landry, N. (2016), '*HoloBot*'. [online] Available at: <https://github.com/ActiveNick/HoloBot>
- Lippert, T. M. (1990), *Fundamental monocular/binocular HMD human factors*, SPIE Proceedings, SPIE, pp. 185–191.
- Lock, E., Kappala, K., Meginness, A. and Shell, S. (2016), '*Sous Chef Demo - Slalom HoloLens Hackathon*'. [online] Available at: [https://www.youtube.com/watch?v=WRkjJ9A3\\_zQ](https://www.youtube.com/watch?v=WRkjJ9A3_zQ)
- Magic Leap (2018). [online] Available at: <https://www.magicleap.com/> [Accessed 15 May 2018].
- Mayring, P. (n.d.), '*Qualitative Inhaltsanalyse*'. [online] Available at: <http://qualitative-research.net/fqs/fqs-d/2-00inhalt-d.htm> [Accessed 14 Apr 2018].
- Mazuryk, T. and Gervautz, M. (1999), '*Virtual Reality - History, Applications, Technology and Future*'.
- Meta Company (2018). [online] Available at: <https://meta-eu.myshopify.com/products/meta2> [Accessed 14 May 2018].
- Meyer, R. (2015), '*Even Early Focus Groups Hated Clippy*'. [online] Available at: <https://www.theatlantic.com/technology/archive/2015/06/clippy-the-microsoft-office-assistant-is-the-patriarchys-fault/396653/> [Accessed 26 Jan 2018].
- Microsoft (n.d.), '*Intelligent bot creation that scales on demand*'. [online] Available at: <https://dev.botframework.com/> [Accessed 08 Feb 2018].
- Milgram, P. and Kishino, F. (1994), '*A TAXONOMY OF MIXED REALITY VISUAL DISPLAYS*', *IEICE Transactions on Information Systems* E77-D(12).
- Nelson, G., Nicoll, G., Luo, X. and Park, J. J. (2016), '*HOLOCOOKING*'. [online] Available at: <http://roxanneluo.github.io/holocook/index.html> [Accessed 15 Jan 2018].
- Niantic, I. (2017), '*POKÉMON GO FEIERT EINJÄHRIGES JUBILÄUM MIT SOMMEREVENTS IN DER REALEN WELT UND INNERHALB DER APP*'. [online] Available at: <https://nianticlabs.com/press/2017/anniversary2017/> [Accessed 20 Feb 2018].
- Orsini, A., Venkatesan, G., Huang, G., Shah, G. and Shah, N. (2017), '*System Overview Augmented Reality Enhanced Cooking With Microsoft HoloLens*'. [online] Available at: <http://www.ece.rutgers.edu/sites/default/files/capstone/capstone2017/posters/S17-35-poster.pdf> [Accessed 13 Jan 2018].

- Oxford Dictionary (2018), '*Definition of hologram in English*'. [online] Available at: <https://en.oxforddictionaries.com/definition/hologram> [Accessed 14 Feb 2018].
- Pagenkopf, A., Engeln, A., Palm, S., Zeiner, K., Burmester, M. and Scheible, J. (n.d.), '*User Experience in SmartKitchen Environment*'. [online] Available at: <http://scheible.hdm-stuttgart.de/smartkitchen/wp-content/uploads/TeaP-Vortrag.pdf>
- Palladino, T. (2017), '*Google Glass Enterprise Edition Open for Business*'. [online] Available at: <https://glass.reality.news/news/google-glass-enterprise-edition-open-for-business-0178875/> [Accessed 19 Jan 2018].
- Papagiannis, H. (2017), *Augmented human: How technology is shaping the new reality*, first edition edn, O'Reilly Media, Sebastopol, CA.
- Park, M. (2017), '*Why Apple Will Win The Augmented Reality Race*'. [online] Available at: <https://www.forbes.com/sites/allabouttherupees/2017/09/07/why-apple-will-win-the-augmented-reality-race/> [Accessed 07 Jan 2018].
- Porter, M. E. and Heppelmann, J. E. (2018), '*Eine Brücke zwischen digitaler und physischer Welt*', *Harvard Business Manager* (2), 22–32.
- Pratley, C. (2004), '*The Acid Usability Test*'. [online] Available at: [https://blogs.msdn.microsoft.com/chris\\_pratley/2004/01/31/the-acid-usability-test/](https://blogs.msdn.microsoft.com/chris_pratley/2004/01/31/the-acid-usability-test/) [Accessed 30 Jan 2018].
- Publikationen* (n.d.). [online] Available at: <https://corporate.vorwerk.de/presse/publikationen/> [Accessed 18 May 2018].
- PwC (2017), '*2017 Global Digital IQ Survey: Augmented reality*'. [online] Available at: <http://usblogs.pwc.com/emerging-technology/2017-digital-iq-ar/> [Accessed 16 Feb 2018].
- Rocket Internet (2017), '*Geschäftsbericht 2016*'. [online] Available at: [https://www.rocket-internet.com/sites/default/files/investors/Rocket%20Internet\\_Annual%20Report%202016\\_German.pdf](https://www.rocket-internet.com/sites/default/files/investors/Rocket%20Internet_Annual%20Report%202016_German.pdf) [Accessed 23 Jan 2018].
- Rolland, J. P. and Fuchs, H. (2000), '*Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization*', *Presence: Teleoperators and Virtual Environments* 9(3), 287–309.
- Rosemary (2018), '*Philip Harben – TV's first celebrity chef*'. [online] Available at: <https://rosemaryandporkbelly.co.uk/philip-harben-chef/> [Accessed 16 Jan 2018].
- Rossen, J. (2017), '*The Tragic Life of Clippy, the World's Most Hated Virtual Assistant*'. [online] Available at: <http://mentalfloss.com/article/504767/tragic-life-clippy-worlds-most-hated-virtual-assistant> [Accessed 29 Jan 2018].
- Rwinj, Zeller, M. and Bray, B. (2018a), '*Gestures*'. [online] Available at: <https://developer.microsoft.com/en-us/windows/mixed-reality/gestures> [Accessed 21 Apr 2018].



- Rwinj, Zeller, M. and Bray, B. (2018b), 'Voice design'. [online] Available at: [https://developer.microsoft.com/en-us/windows/mixed-reality/voice\\_design](https://developer.microsoft.com/en-us/windows/mixed-reality/voice_design) [Accessed 21 Apr 2018].
- Sams, B. (2017), 'Microsoft Accelerates HoloLens V3 Development, Sidesteps V2'. [online] Available at: <https://www.thurrott.com/hardware/90780/microsoft-accelerates-hololens-v3-development-sidesteps-v2>
- Sato, A., Watanabe, K. and Rekimoto, J. (2013), *MimiCook: a cooking assistant system with situated guidance*, in A. Butz, S. Greenberg, S. Bakker, L. Loke and A. de Luca, eds, 'TEI '14', ACM, New York, pp. 121–124.
- Savov, V. (2015), 'Microsoft announces Windows Holographic with HoloLens headset'. [online] Available at: <https://www.theverge.com/2015/1/21/7867593/microsoft-announces-windows-holographic> [Accessed 08 Jan 2018].
- Schart, D. and Tschanz, N. (2018), *Augmented und mixed reality: Für Marketing Medien und Public Relations, 2., überarbeitete und erweiterte auflage* edn, UVK Verlagsgesellschaft, Konstanz and München.
- Scheible, J., Engeln, A., Burmester, M., Zimmermann, G., Keber, T., Schulz, U., Palm, S., Funk, M. and Schaumann, U. (2016), *SMARTKITCHEN Media Enhanced Cooking Environment*, in Unknown, ed., 'IOT2016', ICPS, The Association for Computing Machinery, New York, New York, pp. 169–170.
- Schuhmacher, L. (2018), 'Semi guided cooking app: E-Mail'.
- Sherwin, A. (2014), 'Google Glass To Be Banned From All UK Cinemas'. [online] Available at: <https://www.independent.co.uk/life-style/gadgets-and-tech/news/google-glass-to-be-banned-from-all-uk-cinemas-9570686.html> [Accessed 13 Jan 2018].
- smartKitchen (n.d.), 'Zusammenfassung der Ergebnisse'.
- smartKitchen project (n.d.). [online] Available at: <http://scheible.hdm-stuttgart.de/smartkitchen/projekt/> [Accessed 23 Apr 2018].
- Spangler, T. (2018), 'Magic Leap Rakes in \$461 Million More Funding, Mostly From the Kingdom of Saudi Arabia'. [online] Available at: <http://variety.com/2018/digital/news/magic-leap-461-million-kingdom-saudi-arabia-1202720554/> [Accessed 25 Apr 2018].
- Standefer, R., Iqbal, K., Brandl, K., Dempsey, J. and CashVo, D. (2017), 'Basic features of Form-Flow'. [online] Available at: <https://docs.microsoft.com/en-us/azure/bot-service/dotnet/bot-builder-dotnet-formflow?view=azure-bot-service-3.0> [Accessed 10 Apr 2018].
- Sutherland, I. E. (1968), *A head-mounted three dimensional display*, in Unknown, ed., 'Proceedings of the December 9-11, 1968, fall joint computer conference, part I on - AFIPS '68 (Fall, part I)', ACM Press, New York, New York, USA, p. 757.

- Thermomix – Unser Multitalent* (n.d.). [online] Available at: <https://thermomix.vorwerk.de/thermomix/unsere-multitalent/> [Accessed 16 Jan 2018].
- TK (2016), ‘*Warum kochen Sie nicht?*’. [online] Available at: <https://de.statista.com/statistik/daten/studie/262668/umfrage/gruende-nicht-zu-kochen/> [Accessed 19 Jan 2018].
- Transform your business* (n.d.). [online] Available at: <https://www.microsoft.com/en-us/hololens/commercial-overview#CategoryPivot> [Accessed 12 Apr 2018].
- Turner, A., Zeller, M., Rwinj and Bray, B. (2018), ‘*Gaze*’. [online] Available at: <https://developer.microsoft.com/en-us/windows/mixed-reality/gaze> [Accessed 21 Apr 2018].
- Welzhofer, L. (n.d.), ‘*Was steckt hinter dem Hype um den Thermomix?*’. [online] Available at: <https://www.stuttgarter-zeitung.de/inhalt.begehrte-kuechenmaschine-was-steckt-hinter-dem-hype-um-den-thermomix.621e9dba-8bb5-4a85-b8da-135625543b52.html> [Accessed 24 Jan 2018].
- Werrlich, S., Nitsche, K. and Notni, G. (2017), *Demand Analysis for an Augmented Reality based Assembly Training*, in Unknown, ed., ‘Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments - PETRA ’17’, ACM Press, New York, New York, USA, pp. 416–422.
- Zeller, M. and Bray, B. (2018), ‘*HoloLens hardware details*’. [online] Available at: <https://docs.microsoft.com/en-us/windows/mixed-reality/hololens-hardware-details> [Accessed 05 Apr 2018].