A context aware cooking assistant

Ioanna Lampraki

Thesis submitted in partial fulfillment of the requirements for the

Masters' of Science degree in Computer Science

University of Crete
University of Joseph Fourier

Thesis Advisor: Prof. *James L. Crowley*

This is to certify that **Ioanna LAMPRAKI** conducted an internship in the INRIA laboratory under the supervision of **James L. Crowley**, in partial fulfilment for the degree of Master of Science.

The thesis is entitled:

A context aware cooking assistant

The thesis has been defended in front of and approved by the following jury on September 1, 2014:

Prof. Francois Faure

Prof. Jérôme Gensel

Dr. Dominique Vaufreydaz

Dr. Fano Ramparany

Prof. James L Crowley

September 1 , 2014

Abstract

Context can be said to be the key for the development of intelligent and adaptable systems. In this work, we developed and evaluated a context aware cooking assistant. The approach taken, is to use a situation model for context awareness. Our situation model is structured along time, activity, location and the person's identity. The functionality and usability of the developed system was evaluated by users. The results showed that the users were successfully assisted by the cooking assistant and their performance was enhanced. We conclude with proposals of future work for the system development.

Résumé

Le contexte peut être considéré comme la clé pour le développement de systémes intelligents et adaptables. Dans ce recherche, nous avons développé et évalué un "context-sensible" assistant de cuisine. L'approche adoptée est d'utiliser un modéle de situation pour la sensibilité au contexte. Notre modéle de situation est structuré en fonction du temps, l'activité, lieu l'identité de la personne. La fonctionnalité et la facilité d'utilisation du systéme mis au point a été évaluée par les utilisateurs. Les résultats ont montré que les utilisateurs ont été aidés avec succés par l'assistant de cuisine et leur performance a été améliorée. Nous concluons avec des propositions de recherches futurs pour le développement du systéme.

Περίληψη

Το κλειδί για την ανάπτυξη έξυπνων και ευπροσάρμοστων συστημάτων είναι η κατανόηση και χρήση του όρου ¨πλαίσιο χρήσης' (context). Η παρούσα εργασία πραγματεύεται την ανάπτυξη και αξιολόγηση ενός συστήματος που έχει γνώση του πλαισίου χρήσης(context aware") και είναι σε θέση να βοηθάει τους χρήστες όταν αυτοί μαγειρεύουν παρακολουθώντας τις ενέργειες τους. Η προσέγγιση μας για τη δημιουργία ενός τέτοιου συστήματος, είναι η χρήση ενός μοντέλου καταστάσεων προκειμένου να ενσωματώσουμε την γνώση του πλαισίου χρήσης στο σύστημα. Το μοντέλο κατάστασεων είναι δομημένο με βάση τις διαστάσεις του χρόνου, της δραστηριότητας, της τοποθεσίας και της ταυτότητας του ατόμου. Η λειτουργικότητα και η χρηστικότητα του αναπτυγμένου συστήματος αξιολογήθηκε με την μέθοδο της αξιολόγησης από χρήστες και τα αποτελέσματα της απέδειξαν οτι οι χρήστες βοηθήκαν από το σύστημα και η απόδοση τους βελτιώθηκε. Η εργασία αυτή ολοκληρώνεται με τη δημιουργία προτάσεων για μελλοντική εργασία για την περαιτέρω ανάπτυξη του συστήματος.

Acknowledgements

I would like to thank my supervisor Mr. James L.Crowley for his valuable advice and support during this project. I would like also to thank my colleague Ankuj Arora for his valuable support. Furthermore, I would like to thank Mr. Dominique Vaufreydaz, Mr. Fabien Badeig, Mr. Émeric Grange and Mr. Rémi Pincent for sharing patiently their work experiences with me. Also I would like to thank Mrs. Raffaella Balzarini for her valuable help in the design of the evaluation experiment. In the end, I would like to thank my family and friends for their overall support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter presents the problem of this study, the development of a context aware cooking assistant and the approach taken. Implementation and the evaluation results of the cooking assistant are also presented. Furthermore, this chapter outlines the structure of this report.

People use interactive services in order to perform simple or highly complex tasks as part of their daily routines. Ongoing increase of human needs and expectations make the development of adaptable services a big challenge for computer science. Communication between humans is rich because they share implicit information based on context. Context can be said to have a key role in the development of effective and accurate systems. Context can supplement intelligent interaction by providing essential information.

Everything observable in the environment can be considered as an entity and any information about the status of the entity or the relations among entities can be considered as context. Human environments are rich but an intelligent system must be able to retrieve information about objects or relations that are relevant to the desired task. "Context aware" or "context awareness" is a term that characterizes the services that are able to retrieve the necessary information from the environment in order to appropriately respond and provide the user with suitable information or services. Context awareness aims at affecting the user's behavior and actions so that the user successfully performs the desired tasks. Situation models can be the core component for context awareness to develop more adaptable systems. Human activities can be modeled as a network of situations, each situation representing the state of the environment. Situation models bestow systems with a prediction mechanism of possible next situations based on the observation and interpretation of the user's environment. Furthermore, the attention of the system can be focused and directed by situation models in order to isolate irrelevant entities and relations and select only the required information.

The representation and use of context for computing has been extensively studied as they can be applied to different fields of computer science. Context aware systems can be used for several different purposes such as medical support or gaming. For example, Alzheimer's patients are not able to remember recent events or what they have to do so a context aware service can be used to support their daily activities by continuously observing their actions. In persuasive gaming, the use of context aware systems aims at transferring information such as the user's location and actions from the real world in the game's world. However, context awareness can also be extended in everyday life. Routine activities that people perform in their houses or at work can be also supported by context aware services in order to improve the user's performance. For example, the setup of a formal dinner consists of a determined sequence of instructions about the essential objects and their position. In this environment, people can easily spend too much time to complete the task and they are prone to errors. The presence of a service which is aware of the instructions and observes the user's activities can help the user to successfully complete the task in a shorter time. The goal of the context aware service in each case is to support the user's activities but the required mechanisms of observation and interpretation are different.

The development of intelligent and adaptable systems requires the ability to predict a set of next possible situations in order to achieve an appropriate reaction. Endsley[1] stated that the prediction of upcoming events is feasible when the entities which are selected and are relevant to the task information can be associated with each other in the situation model. According to this theory, situation awareness can be divided in three levels: perception, comprehension and projection. The level of perception includes the sensing of the required information from the environment. Most of the problems that context aware systems face are because they lack accurate or sufficient perception. Focus of attention directs perception in order to select the needed subset of the available information and isolate information of irrelevant tasks. The level of comprehension refers to the integration of data and the determination of their relevance to the user's goal. The selection and comprehension of the relevant information determine the situation. The level of projection focuses on future events, which can be anticipated, after the the current situation has been analyzed.

Endsley's theory[1] can be used to develop effective context aware systems. In order to use Endsley's theory[1], a situation model which consists of seven computational phases, cyclically executed, is studied and used in this work. The first phase, which is called the "sensing phase" is responsible for retrieving the essential environment information from the system's sensors. The attention phase works as a situation filter to reduce the requirements of working memory by directing the sensing phase to perceive only the required information. The assimilation phase is responsible for the assimilation and interpretation of the relevant information in the situation model. The projection and implication phases are responsible for the prediction of the next possible situations and their outcomes based on the analysis of the current situation. The appropriateness of the system actions is determined in the decision phase. The last phase, which is called the action phase, determines how the system must react in each case.

Context aware services can be used to support user's routine activities. In this work, a context aware cooking assistant was developed, based on the situation model that expresses Endsley[1] theory. Before the implementation and evaluation analysis, the utility and the supported functionality of the cooking assistant are presented. Suppose that someone is in the kitchen and wants to prepare something to eat but they need instructions about the needed ingredients or the sequence of preparation steps. Some years ago, the user would search for the recipe in a cooking book. Recently, cooking books were gradually replaced by computers, smart phones or tablets where the user can search and read a specific recipe. However, this method used to cause some difficulties because instructions can be quite confusing or not sufficiently detailed. Thus, a person has to read the recipe several times to figure out what has already been done and what remains to be done. A context aware cooking assistant can be placed in the kitchen in order to replace the use of a recipe. Furthermore, the cooking assistant can prevent users from making mistakes such as missing ingredients or adding a wrong ingredient.

The cooking assistant is able to continuously observe the user's activities in the kitchen in order to provide them with the relevant instructions. At any stage of the procedure, the assistant knows what the user has already done and what they should subsequently do. In the case that the user is confused about the current or the next step of the recipe, the system informs them about it. A common mistake that users do, even if they are experienced in cooking, is missing one or more ingredients. Cooking assistant is aware of the needed ingredients and what has already been added, so if an ingredient is missing, the user is notified. Another common mistake is to get confused about the ingredients and add a wrong ingredient or perform the action of adding the same ingredient too many times. The cooking assistant alerts the users in the event if these mistakes too. In the ideal case of advanced perception, a cooking assistant will be aware of the amount of each ingredient and can inform the user. At the end of the process, the assistant informs the user about the time that the food should stay in the oven or in the fridge. Additional pieces of information such as oven temperature are also provided.

Hence, activity recognition was necessary for our context-aware service. In order to recognize and track the user's activities we used a kinect sensor and a wearable accelerometer to track the user's skeleton and locate all the objects in the scene. The localization of the objects is performed by color detection based on the values of the HSV system. Apart from the position of the objects in the scene , the case that an object is placed in user's hand must be also determined. From the skeleton tracking, the hand and the wrist position were added in the framework in order to calculate the euclidean distance. In order to recognize if an ingredient was going to be added in the bowl, the hand orientation was necessary. For these reason, a wrist band with a mounted accelerator is used. If the tilt of the hand holding a cup is greater than a certain angle, we assume that an ingredient is going to be added in the bowl.

Each situation in the cooking process can be expressed by time, activity, user and location. The situation model that was used in the development of the cooking assistant is structured along these four primary types of context. Each one of the primary context types has a sub-context according to the specific task which represents the specific information which is required.

After the development of the cooking assistant, a sub-system evaluation and an overall system evaluation were performed. The goal of the sub-system evaluation was to test the accuracy and appropriateness of several perception methods because the effectiveness of the system was highly dependent in perception. The results of color detection showed that the light in the scene and the different reflection properties of materials affect the detection accuracy. Furthermore, the overlapping between the used intervals for HSV values reduce the the number of the detected colors. For hand orientation, a Leap motion controller was proposed but the evaluation results showed that the detection of hand gestures was not accurate. The diameter of the object in user's hand affects the effectiveness of the detection. Consequently, the use of Leap sensor was rejected and hand orientation performed by the wrist band with the mounted accelerator proved to be more accurate.

The purpose of the evaluation was to investigate if the context aware cooking assistant would improve user's cooking performance. For the overall system evaluation, a detailed plan of the experiment was determined in order to perform the experiment in an automatic way with each user and come up with reliable conclusions. The hypothesis of this evaluation experiment was the presence of the cooking assistant would improve user' s performance compared to the scenario that a cooking book is used to provide user with the recipe instructions. In order to prove that the experiment hypothesis was true,7 users were asked to prepare two similar recipes with an interval of 7 days. The execution of the selected recipes is similar but the ingredients are different. The selection of having different recipes for each task and the interval between the tasks was decided in order to erase the factor that the users performance in the second task was improved because they remembered the first task.

The functionality and the usability of the system were evaluated in this experiment. Functionality refers to the cooking assistant's ability to follow the routine, detect the user's errors and make the needed alert. Usability refers to the user experience by evaluating user's efficiency, effectiveness and satisfaction. In order to estimate the effectiveness of the system we counted the rate of the people that completed each task. A task was considered as successfully completed, if the instructions of the recipe were followed and all the ingredients were added in the bowl. For estimating the efficiency of our system we had to measure the time that each user needed to complete the task and compare each pair in order to observe if the cooking assistant actually improved user's performance. In order to explore user's feelings and satisfaction, users had to answer in a questionnaire.

Evaluation results showed that the hypothesis of our experiment was correct. The use of the cooking assistant improved user's cooking performance and all the participants completed successfully the task. On the other hand, 28.6% of participants failed to accomplish successfully the first task. Additionally, time measurements of the two tasks showed that on the second task users needed less time to prepare the recipe than the task with the reading of a recipe. Users expressed their pleasure and positive feelings about the cooking assistant and in their comments they made proposals about functions which can be added and they suggested improvements about the supported functionality.

Chapter 2 presents the objective of this work and related studies. Chapter 3 presents the situation model that was used for the cooking assistant and the analysis of the each computational phase. The implementation and the evaluation of the developed system are presented in chapter 4. In Section 4.1 the architecture and the implementation details of the context aware cooking assistant are demonstrated. Section 4.2 details the evaluation of the system components and the evaluation of the overall system. Finally chapter 5 presents the conclusions and directions for future work.

Chapter 2

A Review of Context Aware Services

This chapter reviews the key ideas from the scientific literature on context aware systems. It includes definitions of some key terms and explanation of fundamental ideas used in this project.

2.1 The Role of Context in Intelligent Action and Interation

The term context refers to implicit information that is necessary to understand communication or perform tasks. As an intelligent species, humans undertake highly complex tasks as part of their daily routine. As social animals, humans carry on a rich dialog in their daily interactions. The use of shared implicit information greatly enhances the efficiency of human action and of human to human communication. Thus context can be said to be key for intelligent action and efficient interaction.

Human environments are very rich. Fundamental limits to attention require humans to limit their perception to the relevant objects and relations necessary to perform a task. Similarly, the wealth of information of human to human communication compared to the available communication bandwidth requires use of a minimum number of words and sentences in communication. Context plays a key role for both intelligent action and human to human interaction by providing essential information implicitly.

2.1.1 Context Aware Systems in Mobile Computing

The representation and use of context for computing has been studied since the beginning of computer science[3]. In most cases, researchers have tended to rely on domain specific examples that are relevant only to their specific subdomain [4].

However, as informatics has progressively invaded human activity, understanding of context as a general concept has become increasingly important.

The term "Context aware" was introduced to mobile computing by Schilit and Theimer [5]. In their work, they defined context as a combination of location, users and the objects that are nearby. In their definition they also considered the changes that may arise in these three aspects. Brown et al [6] studied how these systems could be used in everyday life. In this work, context was determined by location and users as well as variables such as day time or temperature. Ryan et al [7] built a context aware assistant for archaeological sites which was able to provide users with information relevant to their personal interests.

Context is increasing associated with a user's location, environment, identity and time. Dey's definition [8] included physical entities such as users or objects and variables such as date, time, location or orientation. Dey also discussed the role of attention and emotional state of the user as aspects of defining context.

Several researchers have presented context as the situation of the environment. Franklin and Flaschbart [9] present context as the user's situation. Context, according to Brown's definition [10], is relevant to the component elements of user's environment and a system can be effective if is aware of them. Similarly, Ward et al [11] considered context as the combination of the elements that exist nearby the application. In addition, Hull et al [12] defined context by the components of the entire environment, not only user's. More technical approaches exist, such as the one provided by Rodden et al [13] where context was defined as the setting of an application.

Several researchers have proposed that factors such as emotion or conceptual skills should also be considered as elements of context. Dey at al [14] referred to the context as the combination of user's physical and social, emotional or informational features. In similar manner, Pascoe [15] defined context as the combination of the physical and conceptual features of an entity, where an entity is anything that can be observed in the environment. Schilit et al [16] indicated that location of the user, who is with him and the available resources are the most important factors for determining context. However, all these definitions related to specific applications and are difficult to generalize. Dey and Abowd [3] referred to context as any kind of information that can describe the situation of an entity. This definition is in line with the definition of context that we have used in this project.

2.1.2 Context Awareness in Human Computer Interaction

In our everyday life, interactive services are available in order to be used for several different purposes. We are interested to have systems that can adapt to the current circumstances and users needs in order to help them to accomplish successfully the desired tasks. Furthermore, systems should be able to take into account the personal preferences of a user. In order to achieve more advanced human computer interaction, the systems must have knowledge about user's context and able to recognize user's activities. The systems that rely on activity recognition must be able to know what information is needed and what they have to observe in order to select it. The purpose is not to select large amounts of data but to select the appropriate information for each case in order to make interaction easier and more accurate. A vital requirement for context aware systems is the real-time sensing in order to select the needed information and respond immediately and correctly to user's needs.[17] During the development of a system, designers should decide and analyze all the possible conditions under which the application is going to be executed and how the system must react in each case.

While, Schilit and Theimer [5] introduced the term of context aware to describe services that can adapt to user's context, more recently, context awareness acquired the meaning of words such as adaptiveness, environment directed or responsiveness. [3] The use of context awareness in computing can be determined by two different approaches. The first approach refers to services that use user's context to recognize, interpret and react appropriate to the user's environment Pascoe et al [15] and Hull et al [12] or systems that model the user's context for interaction.On the other hand, Dey's [8] definition was limited to the aspects of the system interface.

The second approach refers to systems that can use and adapt their functionality to the user's context. Several studies described that context aware systems are able to change functionality based on the user and environment observation. The first reference about adaption was made by Dey et al [14]. They stated that the automated operation of a system can be based on the use of user's context. Ryan [18] gave a more specific definition that refers to systems which use sensors to select information about the environment and give users the option to select context based to their personal preferences. According to Brown's[19] definition they are services that automatically provide relevant information to the user based on the input of the available sensors. The definition which was used in this project was provided by Dey and Abowd[3] and is similar to Brown's[19] definition. According to Dey and Abowd[3] a system can be characterized as context aware if it has the ability to give user information or services that are relevant to his activities. Context awareness can be also referred as situation awareness.

2.2 Situation Models: A Core Component for Context Aware Systems

Human intelligence includes the ability to interpret the meaning of the objects and other people's behavior. An intelligent service requires a similar abilities. Johnson-Laird[20] proposed situation models as a model for human abilities to interpret objects and actions. For Johnson-Laird, a situation is the set of entities relative to as task, properties of the entities such as position and color, and relations between the entities [17]. A situation model is essentially a state description of the that provides context for a task or a communication [21].

Radansky[22] stated that situation model can be adapted to a large variety of tasks. A situation model expresses the implicit information about entities and their relations that is necessary for the task or the interaction. Each change of the entities or relations state is accompanied by the change of the current situation.

As mentioned before, a context aware system must observe the environment components and the activities that take place in order to define the context of use and supplement user's activities. If someone observes the activities that people perform in their daily life, he could say that people are following a predetermined script [23]. When we perform a specific activity in a regular time we use to do it unconsciously in an automated way. Each human activity can be represented as a network of situations. The ability of a situation model to express a human activity requires the development of an ability to observe human activity.

Crowley et al[17] proposed the use of situation models in order to recognize human activities in order to make context aware systems more adaptable . Context aware systems require the observation of relevant entities and relations in order to predict the next possible situations and react in an suitable manner. By the use of situation models, next possible situations can be predicted based on the environment observation and system can response in an appropriate way. Additionally, the use of situation models offers more advantages. The situation models have a reasoning mechanism so they can limit the problems of wrong or missed data. Furthermore, focus of attention can be directed by the situation model in order to select only the needed information. Focus of attention can isolates irrelevant entities and information in order to avoid the overloading of the working memory. In order to study the mechanism and the advantages of situation models we have constructed and evaluated a context aware cooking assistant.

2.3 A Context Aware Cooking Assistant

The purpose of a context aware system is to supplement routine user's activities in order to help them successfully accomplish the desired task and improve their performance. Suppose that, someone is in the kitchen and wants to follow a recipe, for example to prepare a chicken pie but needs guidance about the ingredients or the recipe instructions. In previous years, people used to search cook books about the desired recipe in order to find out the needed ingredients or the instructions. The next step was to place the book somewhere nearby to read and execute the recipe instructions step by step. On the account of the technology evolution, cook books were gradually replaced by computers, smart phones or tablets where user can search and read the desired recipe.

However, the reading of a recipe can cause some difficulties during cooking. Usually, people have to read the recipe more than once to figure out the already executed instructions and what they have to do next. Cooking instructions can be quite confusing or not as detailed as someone might needs. Furthermore, missing steps or errors can not be prevented. The presence of someone who has the knowledge about the recipe and gives instructions to the person based on the observation of person's activities can illiminate these problems. The objective of this work, is the development and evaluation of a context aware cooking assistant that can help user to prepare the desired dish.

The cooking assistant should be able to continuously observe the user's activities in the kitchen in order to provide them with the relevant instructions. At any stage of the procedure, the assistant knows what the user has already done and what they should subsequently do. In the case that the user is confused about the current or the next step of the recipe, the system informs them about it. A common mistake that users do, even if they are experienced in cooking, is forgetting to add one or more ingredients. The cooking assistant should be aware of the needed ingredients and what has already been added, so that if an ingredient is missing, it can inform the user. Another common mistake is to get confused about the ingredients and add a wrong ingredient or perform the action of adding the same ingredient too many times. The cooking assistant should be able to alert the users in the event if these mistakes. In the ideal case of advanced perception, a cooking assistant should be aware of the amount of each ingredient and could inform the user. At the end of the process, the assistant informs the user about the time that the food should stay in the oven or in the fridge. Additional information such as oven temperature could also provided.

In order to represent the approach of the cooking assistant development, we have employed a scenario of a user who wants to make a chicken pie. This scenario was chosen because the specific recipe consists of a well defined sequence of steps and a subset of them must be repeated.

Chapter 3

Situation model for the cooking assistant

For the development of a context aware cooking assistant, the approach of situation models was used. The situation model of this work is based on Endsley[1] theory. This chapter presents the situation model and the analysis of the computational phases.

3.1 General Description

Endsley et al[1] defines situation awareness as " the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". Endsley[1] proposed a theory about situation awareness which claimed that prediction of upcoming events is feasible if we sense entities that the relevant to the user's task and associate the selected information with the situation model. Situation awareness according to this theory can be divided in three levels: perception, comprehension and projection. The perception level includes the sensing of the required information from the environment. Most of the errors that context aware systems face are because of perception problems. Focus of attention controls the subset of the available information that is required and isolates the irrelevant entities and information. Comprehension level includes the integration of data and the determination of their relevance to user's goal. The selection and comprehension of the relevant information determine the situation. Future events anticipation is based on the analysis of the current situation. User's goals and expectations affect the performed processes.

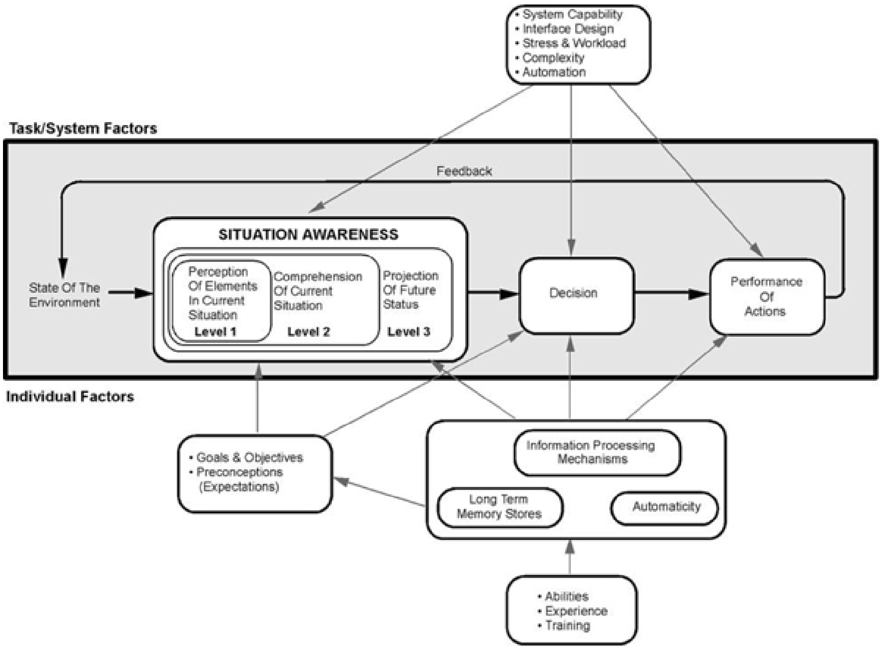


Figure 3.1: Endsley's model for situation awareness[1]

We have adapted Endsley's theory[1] to develop context aware systems. The developed model is represented in Figure 3.2 and consists of seven computational phases that are executed in cyclic way. The sensing phase is responsible for the selection of the environment information from the available sensors. The attention phase guides sensing to perceive only the required information. Attention works as a situation filter in order to reduce the requirements of working memory. The information about the entities, the relations between them and the events in the environment must be assimilated and interpreted in the situation model. The projection and Implication phases can predict the next possible situations and their outcomes based on the analysis of the current situation. The appropriateness of the system actions is determined in the decision phase. The action phase executes the selected action Each one of these computational phases is going to be analytically examined in the following sections of this report.

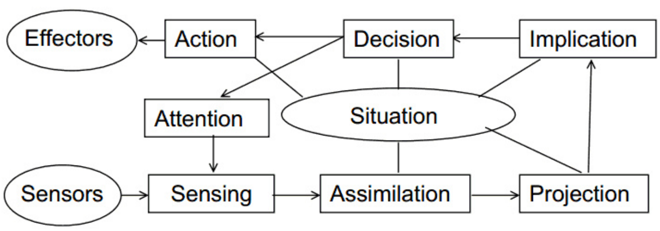


Figure 3.2: Situation Model

The needs or the importance of the perceived information is different for each context aware service. Context must be categorized in order to help designers to encode the possible situations. In this work, situation can be characterized by the types of context which represent the required information. Dey and Abowd[3] claimed that context can be determined and structured along four primary types of context: location, activity, time and the identity of the person.(Figure 3.3) Each one of the primary types of context can be extended by secondary pieces of context that represent the more specific information which must be perceived.

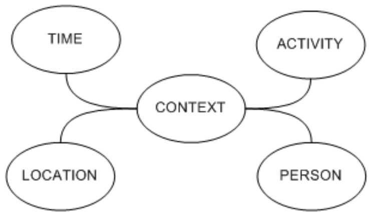


Figure 3.3: Context structure

In order to understand the functionality of the situation model, a scenario that a user prepares a chicken pie is going to be used. Initially, the steps of making a chicken pie can be divided in two groups: stuffing preparation and making the pie. In the beginning, the chicken, the parsley and the peppers must be washed and cut in small pieces with the onions and the smokey turkey. The next step is to put oil in the pan and add the onions and the peppers. They need to be cooked for three minutes and then chicken must be added in order to cook them together for 8 more minutes. The last step is to add the mushrooms. When every ingredient is in the pan, salt and pepper are added and the mixture is cooked for 7 more minutes. The following flow chart represents the steps of the stuffing preparation.(Figure 3.4)

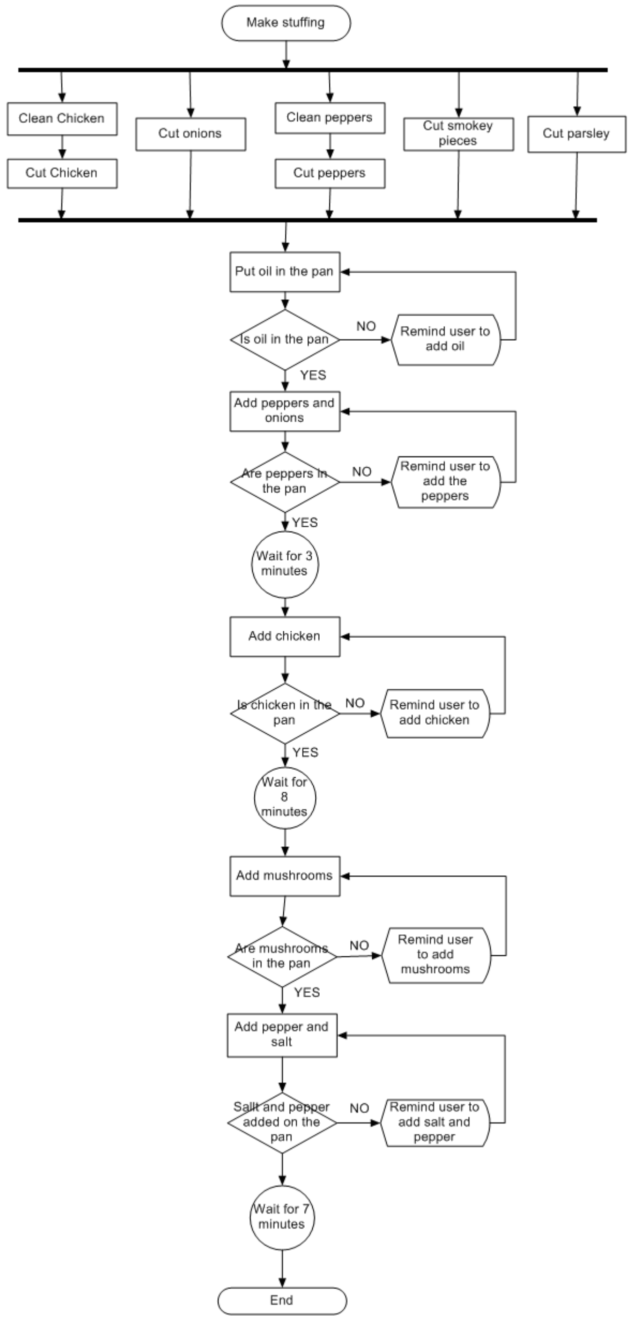


Figure 3.4: Flow chart of making the stuffing of the chicken pie

For making the pie, the first spread must be placed in the pan. In the next step, stuffing, cheese and creme must be added sequentially for 3 times. After the stuffing is ready, we can start making the pie. Initially, we have to put the first spread in the pan. The next step is to add sequentially stuffing, cheese and creme for 3 times. The last step is to put second spread over the stuffing and brush a mix of milk and cheese over the pastry. Before putting the chicken in the oven we have to mix milk with cheese and brush it over the pastry. The sequence of these steps is represented in the following flow chart.(Figure 3.5)

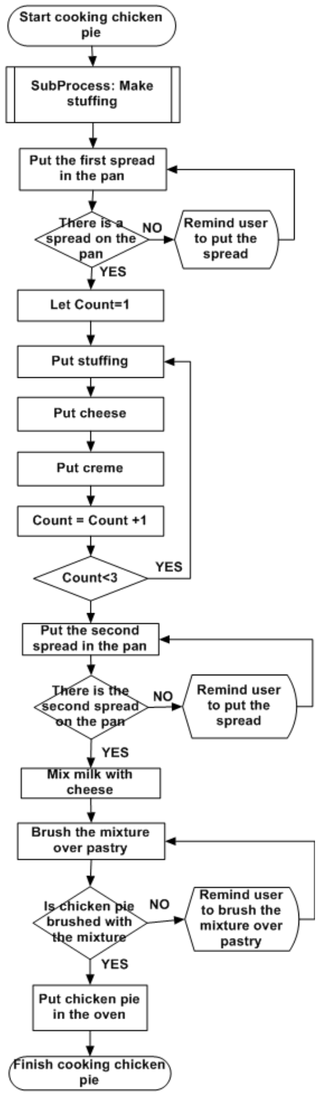


Figure 3.5: Flow chart of making the chicken pie

The situation model for the development of the cooking assistant can be structured along the dimensions of time, activity, user and location.[3] User's context can be described as a graph of the relevant entities and relations.(Figure 3.6) Each dimension defines a sub-context of the developed system. Location refers to the position of the user and the available objects in the scene. Time context is defined by the needed time to prepare the dish or the required ingredients. The available food or material commodities that user must use define the dimension of the person. Activity sub-context refers to the possible actions that user might perform and the instructions of the recipe. Each computation phase of the developed situation model is described in the next section.

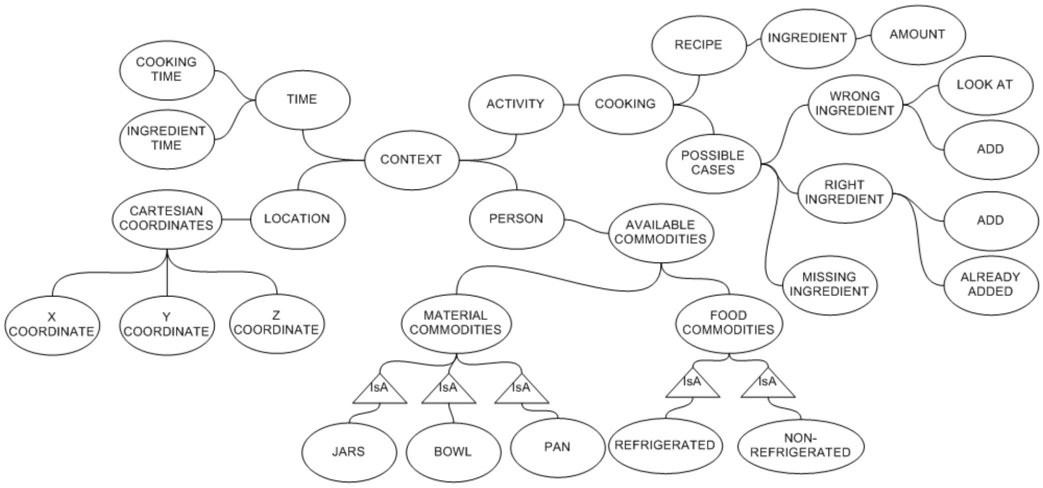


Figure 3.6: Ontology of context aware cooking assistant

3.2 Computational phases

In this section, the computational phases of the situation model are represented. The approach of each phase is represented by analyzing the scenario of a user who makes a chicken pie.

3.2.1 Sensing

The goal of a context aware system is to help users to accomplish successfully the desired tasks. Sensing environment is vital for context awareness cause is one of the prerequisites to specify user's context. The system can not determine user's context or can form a wrong impression about the specific situation if it doesn't have enough information about the environment and the available entities. Furthermore, if the system is not able to achieve real-time sensing, it will not be able to correspond rapidly to user's needs and consequently to be effective.

Sensors are responsible for provide system with information about the environment by making measurements. Nowadays, several types of sensors which can observe/measure a variety of environment variables are available. What must be observed or which sensors are going to be used should be carefully determined based on the needs of the system. It is possible that a service may use more than one data sources which can cause sensing conflicts that should be also handled[24]

In the cooking scenario, the hand of the user and the objects in the scene are the entities that must be observed to select the required information about the task. Real-time tracking of the hand's movement is vital for the recognition of user's actions. Although, user's actions can not be determined without information about the available objects and their position in the scene. User's actions can be specified by the selection and combination of the needed information about the hand and the objects position. When there is no object in user's hand, information about all the objects that are close to the hand must be retrieved. Each object has a specific color, so we have to detect each color that exists nearby the hand and for each one we get the XYZ position. In the case that a user is holding an object in his hand, the system's attention is focused on the specific object and we don't need to sense information about the other objects until the object is placed again on the table. Attention and how is directed is going to be particularly described in section 3.2.7. A context-aware service must be aware of the user's goal in order to help him accomplish the desired task. When the user wants to prepare a dish, the cooking assistant service must know what the user wants to cook and be aware of the specific recipe. The recipe must be read just once in the beginning.

3.2.2 Assimilation

More than perceiving or attending to information, a situation model includes the integration of the information to determine the relevance to person's goal. In the case of the cooking assistant, the position of the hand and the description of the available objects including must be assimilated and interpreted in the situation model. The color In the case that a user holds an object in his hand, only the information related to this object is going to be intepreted in the model.

3.2.3 Projection and Implication

Projection and implication phases predict all the possible next situations and their outcomes based on the current observations and the reasoning mechanism of the situation model. For example, the cooking assistant is able to notice that salt and pepper are going to be added for a second time and consequently the food is going to be salty. In the case that user is going to leave chicken 2 more minutes in the oven, system can predicts that the food is going to be burnt. Another possible case is when the system observes that the user is going to add a wrong ingredient so the recipe is going to be destroyed.

3.2.4 Decision

The decision phase determines how the system should react based on the observation and interpretation of each situation. The context aware cooking assistant makes decision by the analysis technique of decision trees. The approach of decision trees is appropriate for our system because users activities during cooking are sequential and each step is dependent on the previous one. For example, if the first spread of pastry is in the pan, we can add the stuffing otherwise we should add the first spread before we continue with the next steps.

3.2.5 Action

Action phase determines the appropriate system's response for each possible situation. The goal of action phase is to influence users behavior in order to accomplish successfully the desired tasks. The position where the cooking assistant is going to be placed, should be carefully determine in order to be able to observe users actions and provide them with relevant information. In the case that the user is confused about the current or the next step of the recipe, the system informs them about it. A common mistake that users do, even if they are experienced in cooking, is missing one or more ingredients. Cooking assistant is aware of the needed ingredients and what has already been added, so if an ingredient is missing, the user is notified. Another common mistake is to get confused about the ingredients and add a wrong ingredient or perform the action of adding the same ingredient too many times. The cooking assistant alerts the users in the event if these mistakes too. In the ideal case of advanced perception, a cooking assistant will be aware of the amount of each ingredient and can inform the user. At the end of the process, the assistant informs the user about the time that the food should stay in the oven or in the fridge. Additional pieces of information such as oven temperature are also provided.

3.2.6 Attention

Having knowledge about everything in environment is the simplest definition that anybody can use for context awareness. Although, it's difficult for the users to process or combine the huge amount of available information. They need to be aware of the information that is relevant to the goal or task that they want to accomplish. For each task or goal the needed information differ and a good situation model should be able to distinguish and select it from the available information. As we know, on computer science the amount of the selected data is not equivalent to the selection of more information, that's why attention is important for building a situation model. As many surveys confirmed attention has a crucial role in context awareness. Attention is responsible for giving priority to the available information according to its significance. The way that humans direct their attention is quite complex as it is influenced by several factors such as learned scan methods or information sampling. So, one of the biggest challenges in context awareness is to be right about giving priority to the information. From everyday life, users have evolved mental models that determine what they expect in a specific environment. These expectations are also used for filtering the available information.

The attention phase in our proposed model has two significant aspects: guidance sensing phase for perceiving the needed information and reducing the requirements of working memory. Anytime, we need to be able to control the subset of the available information that is required to be selected from the sensors. The decision about which information is going to be processed or combined with other is also demanding for working memory. Specifically, for making a decision we must combine and interpret different information in order to project them. Neurological studies have proved that humans have limited working memory and this adds more constraints to a situation model.[24]

In the cooking assistant scenario, system retrieves information about all the available objects in the scene. Although, when an object is placed on user's hand, system needs information only for the specific object. When the object is placed back on the table, system starts again to observe all the available entities.

Chapter 4

Evaluation

This chapter presents the description of the cooking assistant's implementation. The procedure and the results of the system evaluation with users were also analyzed in this chapter.

4.1 System development

The objective of this project was to build build a cooking service that could help users to cook by giving them information about the desired recipe. In fact, users would be assisted by a robot that was able to recognize their actions in the kitchen. The robot had also knowlegde about the recipe and the available ingredients in order to be able to prevent user from adding a wrong ingredient or a right ingredient in a wrong order. In the case that a user is going to add an ingredient for a second time, the assistant would informed him about.

4.1.1 Image processing for human activity recognition

The cooking assistant's ability to recognize and track user's activities was mandatory for our context- aware service. For this reason, we had to track user's skeleton and locate all the objects in the scene. For the localization of the objects we worked on the RGB and depth images captured by the kinect sensors that are incorporated in the robot. The RGB image (Figure 4.1) was captured directly by the RGB camera of the kinect sensor.

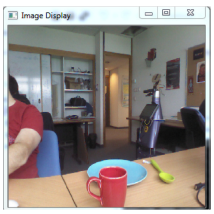


Figure 4.1: RGB image captured by kinect

Kinect also features a depth sensor for capturing depth images. The value of each pixel in the captured image (Figure 4.2) represents the distance between the camera plane and the nearest object(in millimeters)

Figure 4.2: Depth image captured by kinect

By processing these images, our final goal was to define the center of mass for each one of the objects in the scene. For our experiment, we had to handle only one object for each color and the number of the needed colors was 7. Initially, we had to detect each of the colors in the image by the following steps:

Convert RGB to HSV: The image that we receive from the kinect sensor is in RGB format. However, the RGB color space is not suitable to estimate the similarity of two colors. To solve this problem and achieve partial illumination invariance in color recognition, we had to move to the HSV color space.

Build a binary image for each detected color : Binary images have pixels with two possible intensity values, usually black or white. For creating the binary image associated to each color from an RGB image, we have to browse the initial image and associate each pixel's value to 0 or 1. For example, if the $pixel_{i,j}$ of the RGB image has value (0,255,0), then in the binary image associated to the green color, the $pixel_{i,j}$ is set to 1, and for the rest of the binary images, this pixel is set to 0. Figure 4.3 represent the binary images for each detected color.

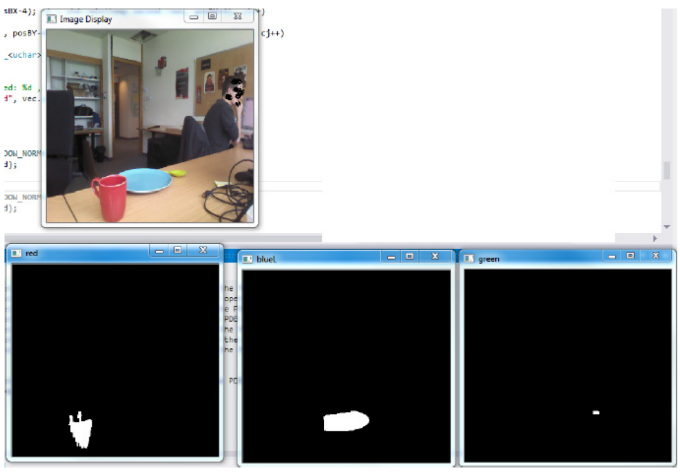


Figure 4.3: Binary images for blue, red and green

To improve system's color detection, we had to make some improvements to the classical algorithm. Initially, we had to make the algorithm able to consider several HSV values for the same colorful object. Thus, if the object has gradual variation in color, each color variation is considered to compute the center of mass of the color object. The principle is to merge several binary images before computing the center of mass. The next improvement was about handling the fact that an object color is not uniform. For this, we had to define an inferior and a superior bound to build the binary image when we search HSV value in the image. This operation consists of run the following calculation with the threshold to 9.

The last improvement was to build a progressive binary image by saving for each color the last 5 binary images. Then we had to apply an updating coefficient of the binary image, to consider a pixel as white we should have at least 3 out of 5 frames of the pixel with the white value.

Compute the center of mass of each white stain in the binary image: This processing step consists of browsing each binary image to compute the center of mass of the white stain. Furthermore, from the RGB image we can detect the position of each color object. The kinect API gives several method to convert the position in the image space to the kinect space and apply euclidean distance to estimate the proximity between the objects. Figure 4.4 shows the center of mass of each detected object in the RGB image.

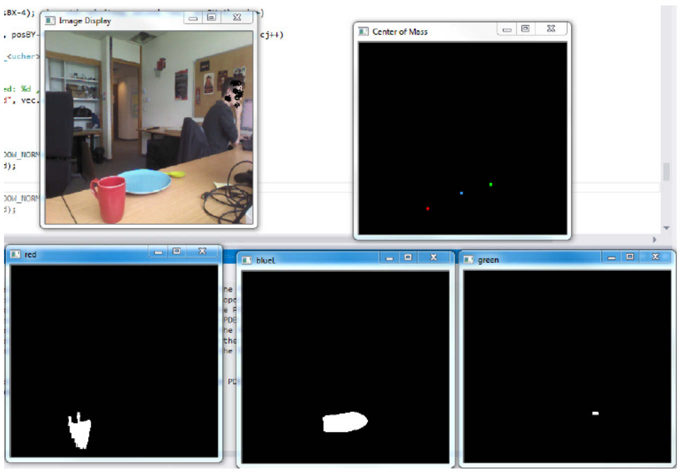


Figure 4.4: Center of mass for blue, red and green object

Except from the localization of the objects in the scene , we had to determine if the user is holding any of these in his hand. From the kinect feature of skeleton tracking, we were able to add the hand and wrist position in the framework in order to calculate the euclidean distance. Figure 4.5 depicts the use of the skeleton tracking when an object is handled.

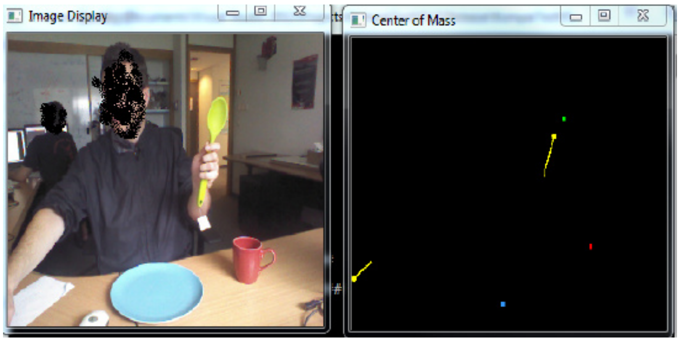


Figure 4.5: Skeleton integration

4.1.2 Leap motion controller for hand orientation detection

By the color detection and the skeleton tracking we were able to recognize the position of the objects and the user's hand but this was not enough. Even if we could perceive if the user was holding one of the objects, we could not know if the user is just holding the object or he wants to add the content in the bowl. In order to gain this knowledge, we needed information about the user's hand orientation. In order to achieve that, we tried to use Leap Motion device (Figure 4.6) that can track hand and fingers movement if the hand is placed over the sensor. This sensor consists of infrared LEDs and two cameras under a black glass. [2]. Our system was going to consist of several components so we had to test separately each one of them in order to ensure the validity of the system evaluation. According to this test, Leap sension was not accurate enough for our task so we had to find another solution.

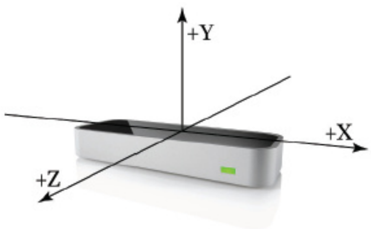


Figure 4.6: The Leap motion controller [2]

4.1.3 Accelerator bracelet for hand orientation detection

In order to be able to detect the orientation of the user's hand, we had to create a wrist band with a mounted accelerator.(Figure 4.7) During the experiment, the user was going to wear it in the hand that he is using for picking up objects. The accelerator is able to determine the angle of tilt of the hand which holds the cup. If the tilt of the hand holding a cup is greater than a certain angle, we assume that the ingredient was going to be added in the bowl. This is done by the accelerator sending a signal to the mini-PC,(Figure 4.8) which then perceives this signal as a confirmation that the angle has crossed the threshold.

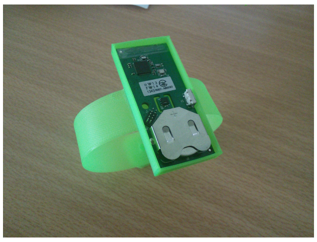


Figure 4.7: Accelerometer bracelet

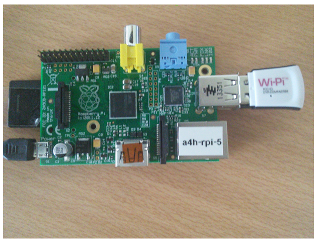


Figure 4.8: Accelerometer mini PC

4.1.4 Cooking assistant android application

The last step of the system development was to build an android application that was running on a tablet that was mounted on the robot. This application is an integrator functionality and acts as the final medium of communication with the user.Specifically, it is connected with the incorporated kinect sensor of the robot and an accelerator mini PC for acquiring the needed information about user's activity. All the components are connected by a custom created WiFi network hosted on a PC and each one of these components is assigned a static IP address. All kinect applications have the feature to behave as a server and open a port for listening for incoming connections. So, all the other components have to connect to the server at the port that he opens.

The information about the hand coordinates, object coordinates and whether the object is held or not is conveyed to the application by means of a socket connection which relays continuously spatial information about the cups and the hand position. From the accelererometer mini pc the android application receives the confirmation message that was sent by the bracelet in user's wrist. In addition, the kinect sensor perceives activity in the scene in terms of movement of the colored cups. The cups are picked up by the kinect and their coordinates in 3D space are recorded. For each cup except from the XY and depth coordinates, we have a variable to keep information about if the object is held or not. All these information are relayed through the opened socket in a JSON format.

Now, we need to have a look at how the android service stores all this information.The app consists of an ontology built in OWL/XML format that stores information about the ingredients of the recipe. Each color is associated to an ingredient entity in which information about the colors and ingredients is stored in the form of pairs. Furthermore, each ingredient has some other properties such as validity, quantity and instruction. Validity value represents if the ingredient is correct or not for the specific recipe. If the ingredient is correct, the quantity property shows the amount that must be added. The instruction property represents how and when the ingredient must be added. At the beginning when the application is loaded, this ontology information is retrieved by means of sparql queries.

Each glass when picked up by the hand is identified by color and this information is relayed to the android app. This color is matched with the color-instruction pairing stored in the ontology and the appropriate instruction is conveyed to the user as a voice command. The android application is responsible for filtering the received information to focus it's attention on the object which is held in the hand of the user. This measurement is done by calculating the euclidean distance between the center of masses of the cup and the hand to compare it to a threshold value. If the distance is below that threshold, we say that the object is held in hand. The accelerometer then proceeds to determine if the held ingredient is going to be added in the bowl or not. This is done by determining the angle of tilt of the hand which holds the cup. If the angle crosses a certain threshold, the ingredient is said to be added. The accelerometer then gauges if the ingredient was added and if positive, relays a confirmation to the android app. This updates the ontology to remember that the ingredient was added and if the user proceeds to add it again, it warns him about the same.

The application gives instructions to the user by voice commands when he picks up a specific cup. The information that is provided includes the name of the ingredient that is inside the cup and instructions about its usage that has four possible cases. In the first case, the ingredient that user picks up is mentioned in the recipe and should be added in this stage.So, the application will say what ingredient is it and the quantity that should be added. Next case is when a user picks up an ingredient that is already added and the service conveys the message that should not be added again. Another possible scenario is that a user decides to add an ingredient in a different order than required by the recipe in this case application mentions that this ingredient should be in the following steps. The last and the most destructive case is when a user is going to add a wrong ingredient and the assistant must inform the user of the error.

4.2 System evaluation

The effectiveness of a context-aware system can be tested by system evaluation. The first approach of system evaluation is the formative evaluation which refers to the evaluation that is taking place during the development in order to improve the design phase. The second approach is summative evaluation that is done after the development in order to investigate if the system fulfills a certain group of predetermined criteria[25]. The goal of evaluation is the improvement of the system by investigating the advantages and disadvantages of our system and identifying problems or deficiencies from the design and implementation phase. In order to achieve our goal, we try to simulate real events by asking users to carry out predetermined scenarios of use. We can determine the needed steps for accomplishing each task and then measure baseline metrics such as time or user's errors. Furthermore, user-centered evaluation is preferred because we can estimate if the time and the form of the delivered information by the context-aware service was useful for performing the tasks. Distribution of the information refers to how the relevant to the task information is requested. There is a large amount of accessible information in the environment that is too difficult for the user to sort it out and filter it. In consequence, the factors that may influence the effectiveness of the provided information is the distribution, the volume and the level of the information.The human factor is imponderable so we have to follow carefully the evaluation procedures in order to end up with accurate conclusions.

4.2.1 Sub-system evaluation

The effectiveness of our system was dependent in perception so before the overall system evaluation, we had to accomplish sub-system evaluation [26] in order to evaluate several perception methods and decide about the appropriateness of each one. In the following sections, the evaluation results of each proposed method are represented and analyzed.

Image color detection and skeleton tracking

The colors that humans can recognize is the result of the light that reflects on the objects surface. Consequently, color detection is dependent on the existed light and the material of the objects. During the development of the color detection, we observed that if we change the light in the scene the results of the detection were different and not accurate so we had to keep the same lightning conditions in the experiment as during the development. The material of the object was also important because light is not reflected in the same way in different materials. For example, colorful glass cups were not detected as good as solid cups with the same colors. Another issue that we had to test was about the number of the colors that we could recognize. On account of that colors can not be represented by a specific number of the HSV system we had to determine a range about the hue value.

This caused problems in recognition of relevant colors because of the overlapping between the value intervals. For example, pink or red objects sometimes were both detected as red. In order to avoid these problems, we had to carefully determine the colors that we were going to detect.

The accuracy of skeleton tracking by the kinect sensor is depending on the position of the sensor in the scene and the observed poses of the human body. [27] In order to limit the amount of data that we had to process, we asked from the user to be sitting behind a table so the kinect could be able to track only the upper part of the body. Furthemore, we tested different positions for the sensor and we conclude that the kinect should be placed in front of the user with a slight downward angle to have a clear view of the table and the user. Except from the skeleton tracking, our experiment was strongly relying on the information about the object held in the hand. As tests show, the detection of the object in hand was an approximation and poor, which means that the appropriate information about the ingredient could not be selected and the ontology could not be triggered. Also since the ingredient could not be narrowed down upon, even if the hand is turned and the ingredient added, the exact ingredient which has been added could not be determined. This would make the experiment very volatile so we decided to switch to a Wizard Oz Experiment.

Leap motion controller

Hand orientation is important for the activity recognition so we had to evaluate if the Leap motion controller was able to provide us with accurate data. On the first task we we tested if the sensor was able to detect hand gestures when the user was not holding an object. The sensor was able to accurately detect every gesture even if the hand was moving quickly. After that, we had to test the case that an object was in the user's hand and the results of this test were not as accurate as on the first test. When the user was holding a cup with a broad diameter such as the hand cannot close around it, the detection was very poor. As Figure 4.9 and Figure 4.10 depict the hand gesture was not always detected.

Figure 4.9: The hand gesture was detected by the Leap motion controller

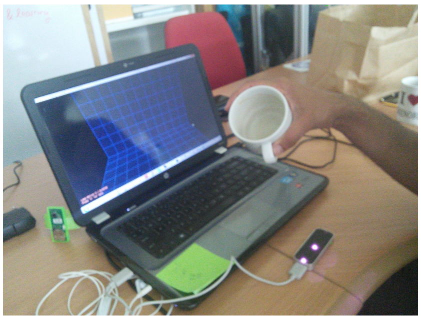


Figure 4.10: The hand gesture was not detected by the Leap motion controller

In the case that the object has a narrow diameter such that the hand can be closed around it, the gestures detection was improved but not completely accurate. For that reason, the method of the Leap motion controlled was rejected. For example if the user was going to add a wrong ingredient but the hand orientation was not detected, the cooking assistant would not be able to prevent the user from making a mistake. So uncertain detection could affect the validity of the evaluation results.

Accelerometer bracelet

After the Leap motion controller was found to be unusable for use in our experiment, we proposed the use of an accelerometer bracelet that could accurately detect a pouring action. During the development of the bracelet, we had to perform several tests in order to specify the angle after which we can consider that the hand is turning. After that, we had to test each one of the tasks that mentioned before for Leap motion controller. The bracelet's detection was found to be completely accurate for our task.

4.2.2 Overall system evaluation

As soon as the evaluation of the separated components was finished, we had to proceed with the evaluation of the whole system[26]. Initially, we had to decide about the purpose of this experiment and specify what we need to observe. After that, we had to make a detailed plan of the procedure in order to perform the experiment in an automatic way with each user. Finally, the last step was the execution of the experiment which include the selection and the analysis of the needed data.

Experiment groundwork

The purpose of this experiment was to investigate if the presence of the cooking assistant would improve a user's performance. So initially we had to decide about some procedural issues such as what recipe we would ask users to prepare. Because of the available perception methods we had to find a recipe with a specific predetermined sequence of simple steps that doesn't need too much time to prepare. Our initial plan, was to ask users to make a recipe from a book and then to ask them to make the same recipe with the help of the cooking assistant. In both tasks we would measure time to estimate if user's performance was actually improved or not. However, the results wouldn't be representative because the second time the user would be aware of the recipe and the improvement of his performance might be influenced not only of the assistant's presence but also because of practicing. For this reason, we decided that we should ask users to make two recipes with similar execution but with different ingredients. In order to eliminate the possibility of a user remembering the first recipe, we decided to perform these two tasks with an interval time of a week.

After that, we had to decide about the number of the participants and determine the characteristics that they should have. We decided that we should ask 7 users to accomplish both of these tasks. In that way, the amount of the selected data was going to be easily analyzed but also enough to end up with safe conclusions. Specific skills or characteristics of the participants where not required for this experiment. However, we decided to have participants that are not very experienced at cooking and they don't know the specific recipes because we assumed that as in real life if someone knows a recipe he would not need help to prepare the specific dish. After identifying the participants, we had to invite people to voluntarily participate in this experiment. When users agreed to participate, an email was sent to each participant to inform about the time and location of each task and confirm interest and availability (Figure 4.11). Initially, the evaluation experiment was going to take place at Inria's kitchen but sensors were not accurate because of the illumination. Consequently, we decided to perform the experiment at the smart room at the smart spaces laboratory where we could control environment conditions such as illumination.

Dear [redacted],

Thank you for your willingness to participate in our experiment. As mentioned before, the purpose of this experiment is the evaluation of the model that we developed for enabling technology for context aware services. You were asked to cook two similar recipes with an interval of 7 days. All the collected data of the procedure (questionnaire's/discussion's answers, video recordings) will be kept anonymous and only the conductors of the experiment will have access to them. The date, time and place of each task are listed below.

First task: Cooking a cheese pie

DATE August 12, 2014
TIME 15:30
PLACE Inria's kitchen close to the reception

Second task: Cooking a courgette pie

DATE August 19, 2014
TIME 16:00
PLACE F103-105

If you need any further information or will not be able to attend for any reason please send an email to ioanna.lampraki@inria.fr. Otherwise we look forward to seeing you.

Figure 4.11: Confirmation email

Experiment hypothesis

Some years before when a user wanted to cook something for example a courgette pie, he had to search for the recipe in a book and read it during the cooking procedure. This was not actually as easy as someone might believes. Reading the recipe may cause some difficulties because of confusing instructions or the need of reading the recipe again and again. Furthermore, the book doesn't have the ability to prevent you from making a mistake. We assume that with the help of a context-aware cooking assistant ,that advise users about next or wrong steps, users are going to be able to prepare successfully a dish in shorter time rather than with a book. So, the hypothesis that we had to estimate with this experiment was that the presence of the cooking assistant would actually improve user's performance.

Experiment protocol

In order to exclude imponderable factors from our evaluation we had to design carefully each step of the experiment:

Introduction: As soon as the user arrived, we welcomed and thanked him for accepting to participate in our experiment. We started with a small introduction about our system and the procedural details of the evaluation. Initially, we informed the user that we developed a theory for building context aware services and that we implemented a cooking assistant in order to demonstrate and evaluate our proposed model. It was important to explain users that the reason of this experiment was to evaluate our theory and not their cooking skills. We assured them that the results of this procedure were going to be anonymous and the conductors of the experiment would have access to them. We also had to inform them that the experiment was going to be video recorded and ask for their permission. We mentioned that any time that they feel uncomfortable they can stop. As soon as they understood and agreed to these terms, we briefly described what they had to do for the specific task.

First task: For the first task user had to make a cheese pie by reading the recipe (Figure 4.12).

Cheese pie without pastry

For 6-8 pieces:
60 gr bacon into small cubes
250 gr feta cheese
200gr grated cheese
100 ml milk
2 eggs
1-2 tbsp olive oil
3 tbsp toast crumbs with pepper and chopped mint

Instructions
In a large bowl, mix the mashed feta with the grated cheese. Add the bacon. Mix the ingredients well and add the eggs. Gradually add the milk and mix well. Then add the olive oil. At the end put the toast crumbs. When the mixture is ready, pour it in the pan.

Figure 4.12: Recipe of cheese pie

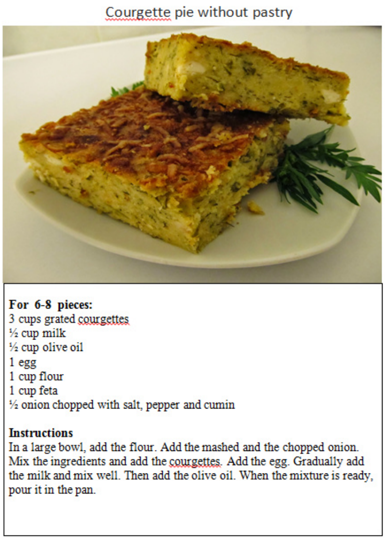
When we finished the introduction, we asked user to take place on the chair behind the table as soon as he felt ready to start cooking. On the table, he could find the available ingredients in colorful jars, a big bowl in front of him for mixing the ingredients and the paper with the cheese pie recipe. In Figure 4.13 we can see the experiment scene for the first task. The whole procedure was video recorded in order to be analyzed after the ending of the experiments. During cooking the user was free to express his thoughts but the conductors of the experiment could not interrupt the procedure or help him.

Brief discussion: When the user finished the cheese pie, we asked him to have a small discussion about this task. We urged him to tell us about the difficulties that he had faced during the cooking and what he would prefer instead of a written recipe. At the end we thanked him for his participation and we remind him the date of the next task. The room had to be set in the initial state before the next user arrived in the room.

Second task: One week was passed before each user had to perform the second

Figure 4.13: Scene of first task

task in the same room. As soon as each participant arrived in the room, we briefly remind him the terms of this experiment. Then we explained him that he had to make a courgette pie (Figure 4.14) for the second task and we asked him to have a quick look at a small paragraph about the recipe.



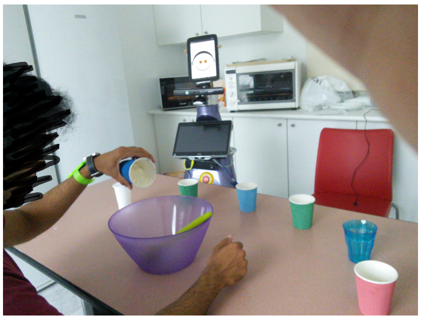
Courgette pie without pastry

For 6-8 pieces:
3 cups grated courgettes
½ cup milk
½ cup olive oil
1 egg
1 cup flour
1 cup feta
½ onion chopped with salt, pepper and cumin

Instructions
In a large bowl, add the flour. Add the mashed and the chopped onion. Mix the ingredients and add the courgettes. Add the egg. Gradually add the milk and mix well. Then add the olive oil. When the mixture is ready, pour it in the pan.

Figure 4.14: Recipe of courgette pie

After he had finished reading, we took back the paper and we asked him to wear the accelerator bracelet and take his place when he is ready to start cooking. Now the user had to wear an accelerator bracelet to help the activity recognition. As the previous task, the user is sitting on a chair with the bowl and the available ingredients in front of him but now he doesn't need the recipe because the cooking assistant is placed on the other side of the table. The whole procedure was also video recorded in order to be analyzed after the ending of the experiments. During cooking the user was free to express his thoughts but the conductors of the experiment could not interrupt the procedure or help him. The cooking assistant observes the actions of the user and helps him when he is missing something or he is going to make a mistake. In the following figures we can see different views of the experiment scene for the second task



Questionnaire: At the end, we asked user to answer an anonymous questionnaire about the user experience of our system (Figure 4.15). When he finished, we accompanied him to the exit and thanked him again about his participation. The room and the equipment had to be ready for the next user.

Task 2: Courgette pie

Please choose the answer that express your opinion by ticking the appropriate box. You are not rated, we are only interested on your personal opinion. At the end you can find space to add your comments.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Unable to Rate
I feel comfortable using this system.	☐	☐	☐	☐	☐	☐
The presence of the robot was disturbing.	☐	☐	☐	☐	☐	☐
I believe I became productive quickly using this system.	☐	☐	☐	☐	☐	☐
I have confidence on the robot	☐	☐	☐	☐	☐	☐
The system clearly informed me about making errors	☐	☐	☐	☐	☐	☐
The system successfully made me avoid errors	☐	☐	☐	☐	☐	☐
The information provided by the system is easy to understand	☐	☐	☐	☐	☐	☐
The information is effective in helping me complete the task	☐	☐	☐	☐	☐	☐
The information was given in an disturbing way	☐	☐	☐	☐	☐	☐
The information is enough for me	☐	☐	☐	☐	☐	☐
Overall, I'm satisfied from the system.	☐	☐	☐	☐	☐	☐

Figure 4.15: Questionnaire about user experience

Evaluation metrics

After the goal and the plan of the experiment, the next thing that we had to determine was the evaluation metrics in order to determine what we need to observe during the procedure. This experiment aimed to test not only the functionality of our system but also it's usability.

Functionality: The supported functionality of a system is determined by a group of rules that developer has defined during the designing and implementation phase. During the evaluation the developer is able to test the functions that system already supports. It's possible that during evaluation developer can realize the need of adding more or removing functions based on users feedback In our case, we had to test if the cooking assistant was able to follow the routine, detect the user's errors and make the needed alerts. For the errors, we had to observe that system was able to detect right the errors and that didn't provide users with false alarms.

Usability: Nowadays, except from functionality, usability is very important factor for the development of a system. Usability of a system refers to the ability of users to successfully accomplish the desired task with effectiveness, efficiency and satisfaction by using this product.

- <u>Effectiveness</u>: refers to the accuracy that users successfully accomplish the desired tasks.
- <u>Efficiency</u>: refers to the relation between the resources and the accuracy and completeness with which users accomplish their desired tasks.
- <u>Satisfaction</u>: reflects the user's feelings and comfort about the use of the system.

Evaluation methodologies

As soon as we determined the evaluation metrics and what we need to observe, we had to decide about which evaluation methodologies were going to be used in order to select the desired data. Evaluation with real users is often used for context aware services because we test if the system was successfully adapted. The purpose of this experiment was to select quantitative and qualitative data by collecting users opinions and observing usage.[28]

In the first task of the experiment, we wanted to identify the problems that the users faced when they had to read a recipe during cooking. In order to achieve that we made a brief interview of the users when they completed the task. In the second task, we chose to use the method of questionnaires in order to collect users opinion about the cooking assistant. The questions were predetermined and same for all the users. Each user had to answer to the questions by selecting one of the provided choices in order to express his agreement or disagreement. In the bottom of the questionnaire, there was space in case that they wanted to add some comments about the system or make proposals about adding or removing features. As the users were going to be asked to use the system as they would do in real life, the experiment was video recorded in order to be analyzed later. Observing the users actions can provide us more information about the user's experience.

Interpretation of results

The purpose of this experiment was to evaluate the functionality and usability factors of the context aware cooking assistant. The whole experiment was successfully set up and working but the approximate information about the detection of the object in hand, could make it volatile and make the conclusions questionable. In order to avoid that we decided to perform a Wizard of Oz experiment [26] where a human is simulating the computer's behavior. For the Wizard of Oz experiment, we followed the procedure of the overall system evaluation. The only difference was that a person was responsible for observing user's actions and providing system information about the object in user's hand. The person used an android application that was running in a mobile. When an object was in user's hand, he had to press the relevant button in order to deliver the relevant information to the cooking assistant android application. As mentioned before, the participants in this experiment were seven.

Functionality: For evaluating the supported functionality of the system we had to measure the times that the cooking assistant provide information to users and observe if the system responses were wrong or right.The following table represents the system responses and how many of them where wrong. As we can see,the system was able to recognize correctly the user's activities and provide them with the relevant information with the few exceptions. Only a few times the system had a false alarm but this was expected because perception can not be 100%. not surprising because perception can not be accurate 100%.

Table 4.1: System responses and errors that the system made

User No	System responses	Number of false responses	Rate of false responses(%)
User 1	9	2	22,2
User 2	14	2	14,2
User 3	15	0	0
User 4	16	0	0
User 5	11	0	0
User 6	12	0	0
User 7	10	1	10

Except from testing the already implemented system functionality, users mentioned in their comments what else they would like of they would except from a cooking assistant service. Several users mentioned that receiving information about the next step as soon as the previous one was completed, was going to be also useful. Furthermore, users mentioned that except from the ingredients, they would like to have more guidance about the recipe instructions such as when and for how long they should mix them. This function was included in our initial design of the system but it was not implemented because of the available perception. As soon as, the perception methods are improved and the need of this function is confirmed by users, it will be implemented as the development of the system proceeds.

Usability: For the system's usability we had to estimate the aspects of effectiveness, efficiency and satisfaction by measuring specific variables.

Effectiveness: In order to estimate if the service effectively helped users to accomplish the desired tasks, we used the rate of the people that completed each task. A task was considered as successfully completed, if the instructions of the recipe were followed and all the ingredients were added in the bowl. In the first task that users had to cook by reading the cheese pie recipe, two of them didn't accomplish the task right because they missed to add one of the ingredients. If we want to express this result with numbers, the 85.71% of the users successfully completed the task and the 14,29% failed.

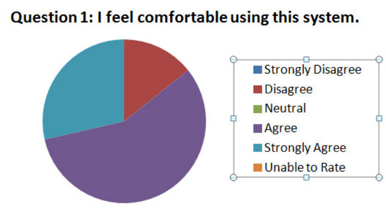
Efficiency: For estimating the efficiency of our system we had to measure the time that each user needed to complete the task and compare each pair in order to observe if the cooking assistant actually improved user's performance. The following table represents all these time measurements.

Table 4.2: Time needed to complete each task(mins)

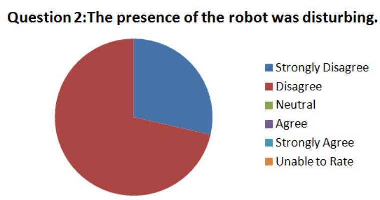
User No	Task 1	Task 2
User 1	3:59	5:14
User 2	8:38	5:32
User 3	5:30	4:21
User 4	6:01	4:32
User 5	5:43	03:56
User 6	4:52	3:21
User 7	7:29	3:48

It is obvious from the comparison of the results that the cooking assistant improved the time that users needed for preparing a dish. The only case that the time was not improved, was the experiment with the first participant where we faced some difficulties with the system set up. What we observed and users confirmed in our brief discussion after the ending of the first task is that they needed a lot of time in reading the recipe because they have to look at it many times in order to find out and understand the next steps. Indicative was the comment of a user that he had to manage many things(read the recipe, understand next step, find ingredient, add ingredient) during the first task so If he had the option he would like to have someone else responsible for at least one of these actions.

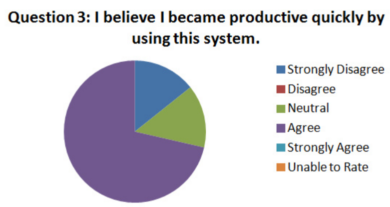
Satisfaction: Except from evaluating if the system works right or improves user's performance, we have to know if the user was feeling pleasant and comfortable during the procedure. Satisfaction factor refers to the emotions that user has about using the system. In order to explore user's feelings we gave him a questionnaire to fill in. The answers of each questions are represented in the following chart pies.



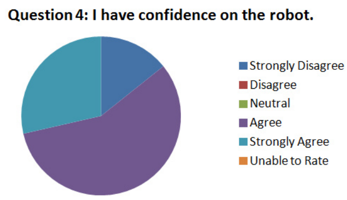
Almost all the users were feeling comfortable about using the system. The difficulties that we faced at the experiment with the first participant obviously influenced his emotions. From the video recordings we could also observe that most of the user were feeling comfortable about interacting with the robot because it was like a game for them.



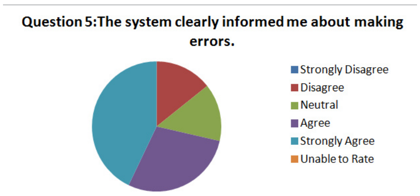
During the development of our system, we concerned if the presence of the robot was going to be weird or disturbing for the users. Even if technology is rapidly developing ,people are not familiar with using robots in our daily life. The results of this questions eliminate these concerns because for all the users the presence of the robot was not disturbing.

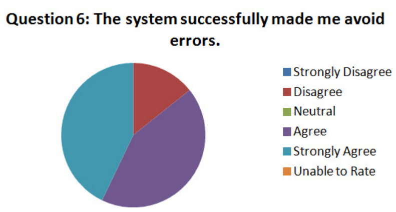


It was important for the evaluation to estimate if the use of the system make users more productive. As the time measurements show, at the most cases the needed time for completing the task was reduced by the use of the cooking assistant.The measurements are representative but it was important to know if the users also felt more productive

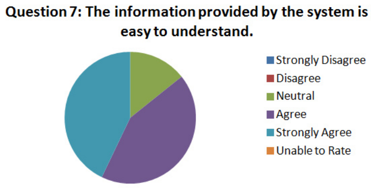


The non disturbing presence doesn't mean that users will follow the instructions they receive from the robot. It was vital for our system that the users feel confidence on the robot in order to follow these instructions. As the results show, every user was feeling sure about the robot instructions and they were willing to follow them.

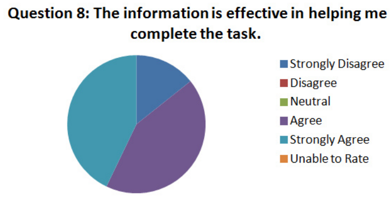




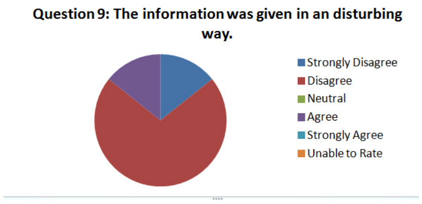
Except from giving the right instructions,the cooking assistant was responsible for preventing users from making an error by providing them with relevant information. The answers of questions 5 and 6 show that users believe that system clearly informed them about potential errors and made them able to avoid making errors.



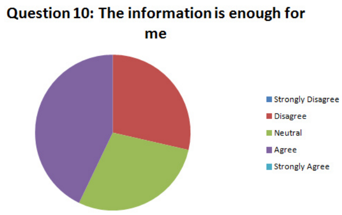
For most of the users was easy to understand the instructions that the assistant provide them because they were short and simple.



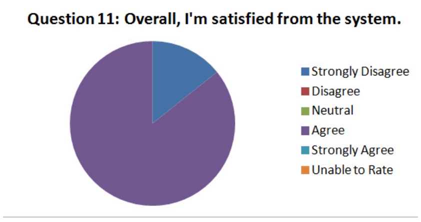
According to users opinion, the provided information was effective for helping them to perform the task successfully.



For the majority of the users, the provided instructions were not disturbing.



Even if the information was effective on helping users and not disturbing, the opinions about the appropriateness of the information amount were divided. Furthermore, users made their personal suggestions about information that may should be included such as more details about the recipe execution.



In the last question we asked users to express if they were satisfied or not by the functionality and use of this system. Most of the users were satisfied but this doesn't mean that we should not make improvements. The evaluation results were going to be used in order to improve our system.

Chapter 5

Conclusions

The use of the cooking assistant improved user's cooking performance and all the participants completed successfully the task. On the other hand, 28.6% of participants failed to accomplish successfully the first task. Additionally, time measurements of the two tasks showed that on the second task users needed less time to prepare the recipe than the task with the reading of a recipe. Users expressed their pleasure and positive feelings about the cooking assistant and in their comments they made proposals about functions which can be added and they suggested improvements about the supported functionality. The proposed cooking assistant has potentials to build a responsive, intelligent system that can be used in real life. This section represents recommendations about the topics that should be subjects for further study. Also, techniques for automatically learning can be a topic for future work. The functionality of the cooking assistant was dependent on human activity recognition. Thus, the accuracy of activity recognition was determinant about the effectiveness of our service. More advanced perception techniques can be studied in order to improve the recognition of human activity. The system used several sensors for retrieving information from the environment and that means that several exchanges of data were performed. Future studies can try to minimize the number of the used sensors for performing accurate activity recognition. Cooking assistant was able to recognize objects by color detection and track the skeleton. Also includes the validation with cooking scenarios that includes more complex activities.

Bibliography

[1] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.

[2] J. C. Coelho and F. J. Verbeek, "Pointing task evaluation of leap motion controller in 3d virtual environment," *Creating the Difference*, p. 78, 2014.

[3] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Handheld and ubiquitous computing.* Springer, 1999, pp. 304–307.

[4] J. Coutaz, J. L. Crowley, S. Dobson, and D. Garlan, "Context is key," *Communications of the ACM*, vol. 48, no. 3, pp. 49–53, 2005.

[5] B. N. Schilit and M. M. Theimer, "Disseminating active map information to mobile hosts," *Network, IEEE*, vol. 8, no. 5, pp. 22–32, 1994.

[6] P. J. Brown, J. D. Bovey, and X. Chen, "Context-aware applications: from the laboratory to the marketplace," *Personal Communications, IEEE*, vol. 4, no. 5, pp. 58–64, 1997.

[7] N. S. Ryan, J. Pascoe, and D. R. Morse, "Enhanced reality fieldwork: the context-aware archaeological assistant," in *Computer applications in archaeology.* Tempus Reparatum, 1998.

[8] A. K. Dey, "Context-aware computing: The cyberdesk project," in *Proceedings of the AAAI 1998 Spring Symposium on Intelligent Environments*, 1998, pp. 51–54.

[9] D. Franklin and J. Flaschbart, "All gadget and no representation makes jack a dull environment," in *Proceedings of the AAAI 1998 Spring Symposium on Intelligent Environments*, 1998, pp. 155–160.

[10] P. J. Brown, "The stick-e document: a framework for creating context-aware applications," *ELECTRONIC PUBLISHING-CHICHESTER-*, vol. 8, pp. 259–272, 1995.

[11] A. Ward, A. Jones, and A. Hopper, "A new location technique for the active office," *Personal Communications, IEEE*, vol. 4, no. 5, pp. 42–47, 1997.

[12] R. Hull, P. Neaves, and J. Bedford-Roberts, "Towards situated computing," in *Wearable Computers, 1997. Digest of Papers., First International Symposium on.* IEEE, 1997, pp. 146–153.

[13] T. Rodden, K. Cheverst, K. Davies, and A. Dix, "Exploiting context in hci design for mobile systems," in *Workshop on human computer interaction with mobile devices.* Citeseer, 1998, pp. 21–22.

[14] A. K. Dey, G. D. Abowd, and A. Wood, "Cyberdesk: A framework for providing self-integrating context-aware services," *Knowledge-Based Systems*, vol. 11, no. 1, pp. 3–13, 1998.

[15] J. Pascoe, "Adding generic contextual capabilities to wearable computers," in *Wearable Computers, 1998. Digest of Papers. Second International Symposium on.* IEEE, 1998, pp. 92–99.

[16] B. Schilit, N. Adams, and R. Want, "Context-aware computing applications," in *Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on.* IEEE, 1994, pp. 85–90.

[17] J. L. Crowley, P. Reignier, and R. Barranquand, "Situation models: A tool for observing and understanding activity," in *Workshop on people detection and tracking, 2009 IEEE international conference on robotics and automation (ICRA), Kobe, Japan*, 2009.

[18] N. Ryan, "Mobile computing in a fieldwork environment: Metadata elements," *Project working document, version 0.2*, 1997.

[19] P. J. Brown, "Triggering information by context," *Personal Technologies*, vol. 2, no. 1, pp. 18–27, 1998.

[20] P. N. Johnson-Laird, *Mental models: Towards a cognitive science of language, inference, and consciousness.* Harvard University Press, 1983, no. 6.

[21] J. L. Crowley, "Context aware observation of human activities," in *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on*, vol. 1. IEEE, 2002, pp. 909–912.

[22] G. A. Radvansky and R. T. Zacks, "The retrieval of situation-specific information," *Cognitive models of memory*, pp. 173–213, 1997.

[23] J. L. Crowley, O. Brdiczka, and P. Reignier, "Learning situation models for understanding activity," in *Proceedings of 5th International Conference on Development and Learning (ICDL)*, 2006.

[24] M. R. Endsley *et al.*, "Theoretical underpinnings of situation awareness: A critical review," *Situation awareness analysis and measurement*, pp. 3–32, 2000.

[25] T. Mantoro, "Metrics evaluation for context-aware computing," in *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia*. ACM, 2009, pp. 574–578.

[26] N. Sanchez-Pi, J. Carbó, and J. M. Molina, "An evaluation method for context–aware systems in u-health," in *Ambient Intelligence-Software and Applications*. Springer, 2012, pp. 219–226.

[27] S. Obdrzalek, G. Kurillo, F. Ofli, R. Bajcsy, E. Seto, H. Jimison, and M. Pavel, "Accuracy and robustness of kinect pose estimation in the context of coaching of elderly population," in *Engineering in medicine and biology society (EMBC), 2012 annual international conference of the IEEE*. IEEE, 2012, pp. 1188–1193.

[28] C. Gena, "Methods and techniques for the evaluation of user-adaptive systems," *The knowledge engineering review*, vol. 20, no. 01, pp. 1–37, 2005.