MIMBCD-UI
Medical Imaging Multimodality Breast Cancer Diagnosis User Interface

Francisco Maria Galamba Ferrari Calisto
francisco.calisto@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Portugal

October 2017

Abstract

Breast cancer is one of the most commonly occurring type of cancer among women, the main strategy to reduce morbidity and also mortality being early detection and treatment based on medical imaging technologies. The current workflow applied in breast cancer diagnosis involves several imaging multimodalities. A need for multi-modal imaging in breast cancer diagnosis is based on the fact that no single modality has the specificity and the sensitivity high enough for reliable diagnosis. Nevertheless, their combination can significantly increase diagnostic accuracy, this also reduces the number of unnecessary biopsies, which leads to better patient care and lower health care costs. It is becoming increasingly apparent in medical image analysis that multiple imaging multi-modalities are required for the accurate treatment and diagnosis of the disease. Our work is to develop techniques that enable the development of an improved user interface breast cancer diagnosis multimodality of image system based on any combination of MG, US, MRI and Text Data. The work involves the development and design of an user interface for automatic detection, segmentation and classification from breast MG, US and MRI, as well as, textual data notations and information visualisation.

Keywords: Human computer interaction (HCI), Usability testing, User interface design, User centered design, Health care information systems, Health informatics

1. Introduction and Motivation

Medical diagnosis for multimodality of imaging is a topic of great interest, it has been the subject of intensive research in the world of medicine. However the developments in terms of breast cancer and innovation in the computational world are still scarce. The Interface herein proposed deals with the analysis of DICOM [17] images taking a multimodality advantage over UI interactions.

Breast cancer is one of the most commonly occurring type of cancer among women [22], the main strategy to reduce morbidity and also mortality being early detection and treatment based on medical imaging technologies. The current workflow applied in breast cancer diagnosis involves several imaging multi-modalities.

A need for multi-modal imaging in breast cancer diagnosis is based on the fact that no single modality has the specificity and the sensitivity high enough for reliable diagnosis [18]. Nevertheless, their combination can significantly increase diagnostic accuracy [2, 14, 19, 24], this also reduces the number of unnecessary biopsies, which leads to better patient care and lower health care costs. It is becoming increasingly apparent in medical image analysis that multiple imaging multi-modalities are required for the accurate treatment and diagnosis of the disease.

Our work is to develop a new UI to collect ground truth (i.e. annotations), concerning two types of lesions that can occur in breast screening, i.e., masses and calcifications. Also, visualizations of the mentioned lesions should be available for a proper diagnosis to perform patient’s follow up.

The task of collecting such annotations is of chief importance. Indeed, with such large annotated datasets it is possible to develop learning methodologies for the automatic diagnosis in the breast screening (e.g. Machine Learning based approaches).
Another important aspect, is that the system should consider several image modalities. This happens, since in current clinical setups [27] the radiologists perform the diagnosis using several image modalities. With such multimodality is possible to obtain a reliable diagnosis, since the complementary information contained in these modalities is crucial for the completion of the exam. More specifically, the following image modalities [13] are considered: (i) CC and MLO Mammography; (ii) US and (iii) MRI volumes.

The problem addressed herein is novel, since the UI should be installed in a clinical domain which constitutes a new paradigm. This work is thus threefold. First, it bridges the gap between the HCI and a clinical domain. Indeed, the literature is still scarce concerning the above regard. Second, it is of crucial importance to know the clinical domain, and to know the workflow of the diagnosis processes that are taken in a radiology room. Finally, the radiologist profile must be known with detail.

Radiologists require rapid and reliable access to images. Within digital environments this access is increasingly based on web browser technologies. Our system specifies a set of web-based services for presenting and accessing a multimodality of medical images. In this system communication the system becomes very flexible. The functionality supports access to a medical imaging data, while client design can be either complex, providing advanced functionalities by taking the medical image data to be applied. All functions (annotations, zoom among others) are therefore provided as controls within the server applications. These may be downloaded locally as pages, so that the only necessary software on radiologist’s workstation is a web browser. At the end, there is no special client application. All the above concerns will be addressed throughout this work.

2. Related Work

In the further section, we discuss software for both clinical and non-clinical tools. More specifically, we address image modality concerning medical applications. Also, we address software concerning the (non)-multimodality views and existing work in the field of medical and clinical user interfaces.

2.1. Medical User Interface

A different interface to visualize patient histories on a PDA [1] describes two different UIs for mobile device tool: (1) one that displays patient histories; and (2) another that permits to visually query patient data stored in hospital database. PHiP (Patient History in Pocket) [1] is a tool designed for a mobile device that displays patient histories and permits to visually query patient data stored in the hospital database exploiting information visualisation techniques where it is able to accommodate on the screen a good amount of clinical cases. The objective of this work is to display as much as possible information about the patient history on a limited display space, providing overview data as well as details. By displaying on a single screen of a personal computer the overview of multiple facets of records will provide users with a better sense of type and volume of available data. We our work, we are going deeper than in the PHiP (Patient History in Pocket) [1] work. In addition of providing a patient (like PHiP [1]) visualization, we also provide temporal visualization of the patients lesion. This work has been developed according to an user-centered approach and will bring system support in that way. Besides the user studies conducted in the hospital at the requirement phase, they have performed with physician evaluations of the different prototypes and this kind of information is kindly useful for our research field as well. While the developed system on that work can provide temporal information of the patient, the authors have optimized the patient visualization. In this work, patient visualization is performed by (i) the temporal morphologic evolution of the masses; and (ii) the temporal evolution of the calcifications. In short, the system has the capacity of merging both patient visualization and patient’s follow-up. A dynamic information over time.

In computer-assisted creation of patient progress notes [28] a prototype application, called activeNotes, supports the creation of critical care notes by physicians in a hospital of an intensive care unit. It integrates automated, context-sensitive patient data retrieval and user control of automated data updates and alerts into the note creation process. A critical care note is a clinical document, written by a hospital physician, that documents the patient’s progress and prognosis. This kind of work will help our project system by giving us information analysis from physicians feedback and user understanding. It also bring us information with qualitative study by providing us the right path throw prototype design and user experience evaluation. The physician-driven management of patient progress notes in an intensive care unit [28], describes the design of an exploratory focused techniques to support data input and management of electronic progress like note content. This will help the project system by giving us an alternative design exploration including observations, structured and semi-structured interviews, design and implementation of a prototype, as well as, feedback gathered in qualitative study with physician surveys. This approach was not yet implemented in the morphologic/density diagnosis of the breast cancer lesion. Something that emerge on this our work.
In many contemporary environment systems, there is a big opportunity [10] to improve the UI. From a collection of displays, complex and tedious procedures, inconsistent sequences of actions, inadequate functionality and insufficient informative feedback can generate debilitating anxiety and stress. In this section it is described the most relevant interaction environments [3], like traditional (mouse and keyboard), touch and virtual reality. Another relevant aspect that deserves interest is the presence of bias. This can be regarded as the consistent differences between the location that users want to interact with, and where they actually interact [11]. Concerning the biases created by radiologists at various positions relative to the device environment, the results found by Lundstrom et al. [15] indicate that device biases depend on viewing angle. In our work, we must overcome the challenges inherent to evaluate multiple variables, noisy experimental results and difficult to measure metrics while advance our knowledge regarding device interactions. Also, quantify existing knowledge. This is an improvement, but still scarce to achieve quantitative results or, even more important, statistical significance about those device interactions in context of clinical domain. Such literature from this authors is of an immense importance research supporting our research understanding and system decisions, by giving us context over the domain.

2.2. Systems

Studies on breast cancer diagnosis system have been approached for many years from different angles of view: the cause of the lesion; the detection of the lesion; the diagnosis of lesion; the diagnosis systems; method of treatments before and after surgery. Two paradigms divides these studies. The first one, which defines breast cancer lesions as a local and regional lesion. The second one, as a systematic lesion control where systems are helpful to diagnose and follow-up patients, as a preventive act and an early detection of lesions. The breast cancer diagnosis systems work [12] is an overview of the studies that have been done to assist medical systems and user interfaces in producing more accurate and faster diagnosis of breast cancer patients.

Systems that focus on data extraction, like the one Ganesan et al. [9] have described, a computer aided system that can do malignancy probability estimation of mammography lesion assisting radiologists to decide patient information management while improving the diagnostic accuracy.

The R2 ICS [26], identifies potential ROI from the detection of clusters of microcalcifications or speculated masses. First, radiologists reads the images, then views the results of the ICS analysis on a display monitor. The radiologist may then return to the original image to confirm whether or not anything was missed at location on the monitor. Initiatives to nuclear medicine imaging and other techniques in breast cancer diagnosis have been taken by researchers to explore the nuclear medicine field for an improvement in the diagnosis of the lesions. This new approach is a novel way of using CADx technology, extending the paradigm of providing radiologists a follow-up feature of the annotated lesions. This system underlines the importance of an ICS, or more commonly called by us as follow-up feature, as an important feature to our system.

A work from Chen et al. [4] claims that a reliable and effective screening mass protocol for patients at risk with age ranged from 40 to 49 years old does not exist. An early pre-menopausal women patient at this age range have a denser breast tissue, therefore any lesions, masses or calcifications are not easily detectable by the mammogram. Contrast-enhanced magnetic resonance imaging, in the recent development, has been indicated as a promising complementary technique to mammography as a potential tool for screening of younger women patients, due to its three dimensional characteristics. The MRI based imaging tool does not require the use of ionizing radiation [29] but it produces higher diagnostic sensitivity, especially in case of dense breast tissue. The MRI imaging single modality tool [21] is a web-based CADx system for manual and automatic extraction and analysis of breast masses. Thus, the radiologists use a normal browser to view scans from a remote station, in order to exchange of know-how among the clinicians and improve data management. The proposed systems are important to our research, while it outcomes the promising results and facts. Therefore the web-based tool is likely to be improved by using accessible architectures for a better radiologist’s performance [20] on the diagnostic relations between curve morphology’s, enhancement values and lesion detection. Emphasising the same need for accessibility in our system.

Realizing the crude drawback of a mammography-based breast diagnose system, several studies [23, 8] are improving our data extraction technique by merging computer aided systems in lesions extraction and progressive images with microcalcification. Different approaches were used in the single-modality to extract data, or to be more precise, to extract female patient breast images with higher resolution, with a goal to produce a system that could detect the smallest tumor [6] in breast cancer patients with less false-positive cases. What leads to unnecessary biopsies. Concentration on detecting tumor in breast with denser tissues (younger pre-menopausal women) is also done in a long aim to reduce the mortality rate due to this disease.
3. Identified Features

To include the radiologist in the design and development of a new medical and clinical information system, we adopt a participatory approach for user requirement gathering and prototype development. Having said that, our experience is described in developing prototypes and aims to report the demonstration of some of the second Low-Fi Prototypes developed in Balsamiq [7]. The quality of data is a product of tools and techniques adopted. Requirement specifications templates, wireframing tools (e.g. Balsamiq Mockups) and frequent iterative knowledge with stakeholders, play a key role in our implementation of the participatory approach. Based on user requirements and paper prototypes, the UI development can be triggered. Balsamiq Mockups [7] are a rapid wire framing tool which can build a rapid prototype in the software engineer, as said before. It can be used to draw an interface sketch for user interaction. Once radiologists find out functional and practical, it can be treated as the High-Fi Prototypes. UI prototyping is the most critical part in this research work. It should follow the UI rules and meet the users need at the same time. The final prototypes of this Low-Fi phase should meet following requirements:

- The UI should provide support to the user, that should understand the multi-modality of imaging interaction. So radiologists can better and faster diagnose breast cancer.
- The UI should help the user find where are the breast masses and calcifications.
- The UI should help young inexperienced radiologists to quickly familiar with user interface.
- The UI must allow the radiologists to visualize the three modalities (MG, US and MRI) in different screen levels and sizes measuring and diagnosing the breast cancer.

Something immediately pointed out by radiologists was the need to compare the last two MG modality images (CC and MLO) acquisitions on a main screen, but it is trivial that this can not be the first screen to appear. However the preferences were made, we must not forget to have a clear option screen as first to choose the patients (PACS) name and days. A proposed second screen will have four of the multi-modality of imaging, and we will consider that as a main screen, since the patient was already chosen. From the requirement of having the set of CC and MLO screen views we implement a prototype with most of the screen directing to this option. On the other hand, it is not enough information to shown on this screen and it is fundamental to have a set of screens with the last two date image acquisitions.

4. Implementation

In a project like this, a major decision to take care about is the technologies choice, like programming languages, libraries and frameworks. Although our experience in JavaScript, HTML and CSS was not at an expert level of the technology, we knew that it was a very flexible set of technologies for a UI multimodality of imaging development, also, allowing a multi-paradigm approach (JavaScript), such as (i) Object-Oriented; (ii) Interpreted; (iii) Client Side; (iv) both Imperative and Declarative; and (v) Functional. Furthermore, it has a very complete standard libraries for our Research Project, allowing the prototypes development.

Visualization tools for the DICOM objects allow configuration and integration with an existing PACS. Among these tools we can find support of Cornerstone [5] libraries. These libraries have in common being based in HTML5 and JavaScript so they allow the manipulation, display and access to DICOM object information completely in a web browser. For the implementation of our MIMB-CDUI project (WADO services), accessed by our UI Client service, part of the DICOM standard, in order to determine the parameters, it must be accepted as a service of this type and must be the response answer of the service. The standard defines parameters in two variants; required and optional. The required parameters are those that identify unequivocally a DICOM object; the study identifiers, series and instance. All other parameters refer to modifications on the image or format of the data. In addition, the standard includes an exception with respect to these parameters. If a PACS server supports search by unique instance identifier, parameters of study and series are not necessary. This is the case of a server, which the MIMB-CD-UI project is implemented in, so the amount of parameters that should be treated and validated when searches are reduced. Figure 1 shows a UML sequence corresponding to with the execution of an HTTP request on the UI (WADO service). The design of the service is explained below. Upon receiving an HTTP request, it collects the parameters and links them to a data model. If the request is validated correctly, the service searches for the location of the object in the PACS and if available responds to the HTTP request by returning the object’s data. Otherwise, the request is answered with a HTTP error code as specified in the standard WADO. Possible answers are 406 Not Acceptable for Invalid Requests and 404 Not Found for searches without results.

This phase of the project marks an important turning point for the development of the same. In this phase it took carried out a research on the different types of DICOM viewers and similar tools that are available. The objective of this research
was to determine the most suitable for the development of the DICOM based viewer as a browser UI, therefore the focus of search were solutions web-based. The solutions researched were mainly from Open Source projects.

5. Evaluation

Formal verification of statistical reactive data systems typically requires checking quantitative and qualitative properties. The former means that checking whether a certain event path (tasks) holds almost surely checking all path, while the task of a quantitative analysis is to check whether a giver (upper or lower) statistical information bounds for a certain event path while it can be guaranteed for all subjects.

5.1. Performance & Experience

Performance and experience are addressed within-subjects design with the UI prototype. Figure 2 shows the order of the user (radiologist) tests. First, the user took a Demographic questionnaire on an online form (Electronic Questionnaires). Second, a user Training session on a mouse-based environment, where the researcher show to the user what are the features and a resume of the task. Third, the user executes a set of pre-defined Tasks. The fourth phase is followed by a questionnaire where the researcher ask the user information about Experience of use. At the end, another Final online form (Electronic Questionnaires) is answer by the user.

The after-task questionnaire concerns the Experience of use (UX) that takes place after each condition test, resorting to a set of questions rated on a Likert-scale. The user is inquired about the following nine conditions:

1. Competence;
2. Autonomy;
3. Relatedness;
4. Immersion;
5. Positive Affect;
6. Negative Affect;
7. Enjoyment;
8. Intuitive Controls;
9. Preferences;

5.2. Results

We analyzed the performance using several metrics including, time, number of interactions, hit rate score and errors. During the test some notes are collected to support results, that includes preferences, improvement suggestions and applicability of the UI prototype performance. Since the data collected does not meet the applicability pre-conditions required for ANOVA, i.e., does not follow a Gaussian distribution and also the number of samples is small (eight radiologists), we resort to use the Kruskal-Wallis Analysis of Variance [25] to test the performance and experience that is useful for testing between-subjects effects for three or more conditions [16].

Table 1: Time Performance Results.
mentioned, for all the events, all coordinates of the annotation are saved and all videos involving the user interaction are recorded. The relation between the radiologist expert level is also important here. The average of each subject’s successful tasks is used to obtain the completion time aggregate by radiologists. The unsuccessful tests were excluded in the completion time analysis. By evaluating the completion time of the task, the analysis revealed \( \text{M}_{\text{time}}=168.38 \text{s}, \sigma_{\text{time}}=64.12, \text{sem}_{\text{time}}(\bar{x}) = 22.6698 \) a radiologist maximum time spent of 296 seconds and radiologist minimum time spent of 103 seconds. The conducted post-hoc tests, by the application of THSD test, revealed that there are differences between all the interactions in terms of completion time.

Table 2: Number of Annotation vs HRS Results.

The number of annotations have a direct relation between the HRS while we compute the percentage annotations (Table 2) that were scored as a score of three points (HRS = 3) or a score of one point (HRS = 1). For instance, as more accurate annotations are done, better the information data will be generated. That way, as more percentage of the red bar we have, what means that the percentage of high score (HRS = 3) is higher. With a deep observation we can conclude that radiologist five was the one who did more annotations and in the same time who did the most accurate (Table 2) ones. The throughput of the loss was the time (Table 1), if we observe the time spent for this radiologist five, we can see that was the slowest one. Typically understandable, while radiologist five did the most number of annotations and the most accurate ones.

From the Table 3 we can observe how the relation between annotations and total number of interactions (\( M_{\text{noi}}=62.38, \sigma_{\text{noi}}=50.177, \text{sem}_{\text{noi}}(\bar{x}) = 17.7402 \)) will behave, where the mean of the number of interactions (\( M_{\text{noi}}=62.38 \)) is 11.62% more (\( M_{\text{noa}}=55.13 \)) than the mean number of annotations. However the observation of several differences between number of annotations (\( M_{\text{noa}}=55.13, \sigma_{\text{noa}}=39.39, \text{sem}_{\text{noa}}(\bar{x}) = 19.49 \)) and number of interactions were noticed, the difference of the means are not so relevant and were typically expected, while the same difference was observed on both \( \sigma_{\text{noa}} \) and \( \text{sem}_{\text{noa}}(\bar{x}) \).

Another number, the error counter, may have been regardless, so the method chosen as error counter did not affect the conclusion about overall measures from performance and the total number of iterations process. The errors must take in consideration when trying to understand accuracy and time spending on annotations repeat. A typical criteria and measures included the length of time it takes typical users to learn to understand the UI to do the schedule tasks, and the number of errors they make, shows that (\( M_{\text{noe}}=1.50, \sigma_{\text{noe}}=1.414, \text{sem}_{\text{noe}}(\bar{x}) = 0.4999 \)) our UI is enough robust from the mean value of 1.50 errors/radiologist number of errors. What means that for each radiologist that take a mean time of interaction (\( M_{\text{time}}=168.38 \text{s} \)) will do a 1.50 error per radiologist. This is completely acceptable and less then expected.

Table 4: Positive and Negative Affect.

The mean significant (\( M_{\text{pa}}=4.38, \sigma_{\text{pa}}=0.518, \text{sem}_{\text{pa}}(\bar{x}) = 0.1831 \)) main effect of Positive Affect (Table 4), while the median value (\( m_{\text{pa}}=4.00 \)) is the same of the minimum value (\( \text{min}_{\text{pa}}=4.00 \)), less than the maximum value (\( \text{max}_{\text{pa}}=5.00 \)). A percentage of 37.50% voted \( x_{\text{pa}} = 5.00 \) that corresponds to the maximum value attributed on this results.
Finally, a percentage of 62.50% voted $x_{na} = 4.00$ what corresponds to the minimum value attributed on this results. Whereas the marginal effect (Table 4) of the Negative Affect ($M_{na}=1.50$, $\sigma_{na}=0.756$, $\text{sem}_{na}(\bar{x}) = 0.2672$) is commonly expected in comparison to the good results of Positive Affect. Also, analyzing the minimum value ($\text{min}_{na}=1.00$) with $62.50\%$ of the radiologists, and a maximum value ($\text{max}_{na}=3.00$). A percentage of $62.50\%$ voted $x_{na} = 1.00$, a percentage of $25\%$ voted $x_{na} = 2.00$ and finally a percentage of $12.50\%$ voted $x_{na} = 3.00$.

### Table 5: Intrinsic Motivation Inventory.

<table>
<thead>
<tr>
<th>IMI</th>
<th>Enjoyment</th>
<th>Competence</th>
<th>Autonomy</th>
<th>Relatedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

The results ($M_{enj}=4.25$, $\sigma_{enj}=0.463$, $\text{sem}_{enj}(\bar{x}) = 0.1636$) from the IMI and more specifically about enjoyment (Table 5) were clearly positive. Where, a high percentage of $75\%$ voted $x_{enj} = 4.00$ and the other $25\%$ voted the maximum value of $x_{enj} = 5.00$ from the 5-point Likert-scale. The median ($m_{enj}=4.00$) of the enjoyment was also average high. Radiologists also perceived themselves to feel competent (Table 5) using our UI prototype features and tools. A percentage of $50\%$ selected the maximum value of $x_{comp} = 5.00$ from the Likert-scale, while the other $50\%$ choose the $x_{comp} = 4.00$. Here the results are even higher ($M_{comp}=4.50$, $\sigma_{comp}=0.535$, $\text{sem}_{comp}(\bar{x}) = 0.1891$). The median value ($m_{comp}=4.50$) represents also an evidence that radiologists felt competent using our UI prototype. Describing the autonomy (Table 5) that in comparison to competence, also showed that our UI prototype gave a good experience to the radiologists related on those conditions (competence and autonomy). The results ($M_{aut}=4.50$, $\sigma_{aut}=0.535$, $\text{sem}_{aut}(\bar{x}) = 0.1891$) showed that, like on the competence condition, a percentage of $50\%$ selected the maximum value of $x_{aut} = 5.00$, while the other $50\%$ choose the $x_{aut} = 4.00$. The median value ($m_{aut}=4.50$) represented also an evidence that radiologists felt autonomous. As a scale, relatedness (Table 5) assesses the affective, cognitive and experimental aspects of radiologists’ connection to an early identical experience. Observing the results we can conclude that a percentage of $50\%$ from the radiologists, felt highly related ($x_{rela} = 4.00$) to our UI prototype. A percentage of $37.50\%$ ($x_{rela} = 5.00$) felt extremely high related to the UI prototype, while only $12.50\%$ ($x_{rela} = 3.00$) felt normally related. The mean value ($M_{rela}=4.25$, $\sigma_{rela}=0.707$, $\text{sem}_{rela}(\bar{x}) = 0.2499$) showed to be higher than average, so relatedness is also a conserved condition to the study. This scale correlated as expected with other condition measures, and relatedness was a better predictor of radiologists involvement, sustainable consumption, and user identification. The relatedness results support the validity of affective, cognitive and experiential aspects of radiologist connection with the UI prototype.

### Table 6: Experience Needs of Satisfaction.

Several aspects related to immersion were discovered in this results that corresponds to the existing research of the immersion scale. Discussing the entry into a diagnostic, assessment is an integral for any level of radiologist engagement with a tool on the radiology room. The more a task is oriented to the UI, the more users will put up with minor usability issues if the overall experience is pleasurable. For instance, our immersion levels showed up a significant positive mean effect ($M_{immme}=4.00$, $\sigma_{immme}=0.756$, $\text{sem}_{immme}(\bar{x}) = 0.2672$) for this condition. While $25\%$ of the radiologists voted $x_{immme} = 5.00$ that represents the maximum value, another $25\%$ voted $x_{immme} = 3.00$. That means a $50\%$ voted $x_{immme} = 4.00$, a value average representing the majority as the median value ($m_{immme}=4.00$) also showed. In contrast, engagement, and therefore enjoyment through immersion, is not possible if there are usability and control problems. Usability flaws could hinder this. Using specific subscales, several hypothesis are tested from the ENS conditions. We first explored for main and interactive effects of radiologists on each task variables. Results revealed to be positive ($M_{ic}=4.75$, $\sigma_{ic}=0.463$, $\text{sem}_{ic}(\bar{x}) = 0.1636$) to this condition. Mainly, $75\%$ of the radiologists said that the controls were highly ($x_{ic} = 5.00$) intuitive, while the other $25\%$ answer the controls ($x_{ic} = 4.00$)
were intuitive. As preferences are differently scaled \( (1 \leq \bar{x}_{\text{pref}} \leq 4.00) \), the mean value \((M_{\text{pref}}=2.63, \sigma_{\text{pref}}=0.916, \text{sem}_{\text{pref}}(\bar{x}) = 0.3238)\) showed to be more average \((m_{\text{pref}}=3.00)\) meaning that radiologists preferred our tool for the purposed solution.

Statistical analysis revealed that expert level of the radiologists had an influence on the performance and experience conditions. Also, the results showed that the UI prototype is solid enough to the pursue and can be put on production. This research is a part of a larger project aimed to develop a multimodal of medical imaging UI for the breast cancer diagnosis allowing radiologists to visualize and annotate several modalities of medical imaging. Future analysis will consider a deeper and more complex results on the performance and experience of the user, as well as validate the present results.

5.3. Discussion

The exploratory research in this section and the performance and experience evaluation of the MIMBCD-UI project in this section yielded interesting points of possible functional improvements. Also, the System Architecture section is important to debate several system decisions and their improvements. During the performance and experience evaluation radiologists were given the possibility to give feedback on the test case system and indicate whether they missed certain functionality and other issues they would like to see improved. These improvements are subdivided into the following categories: (i) functional improvements; and (ii) UI improvements. There are no direct effects between the affect, intrinsic motivations and needs of satisfaction results on any of the experience measures. The only interaction involving both performance and experience suggests that each condition has its own pair of relations. Furthermore, this results show that radiologists are significantly happier (positive affect) remaining on their usual work-flow tool. In addition, radiologists have higher competence, autonomy, relatedness and immersed, with a mental involvement on the tasks, by the use of our UI prototype.

Functional improvements for the MIMBCD-UI project include the option to switch between patient studies. This is very important because radiologists often want to look at different studies or old and new image studies of the patient. Also it should be possible to filter between image source type, filtering between CT, MR and X-ray scans for example. Ideally MIMBCD-UI project should also be able to display 3D image reconstructions and allow for 3D interaction, but this is not a must and would require a more complex interaction style in three dimensions. Less frequently mentioned but nonetheless possible valuable suggestions include the addition of optional patient information. The MIMBCD-UI project UI can be improved by extending the current interface with a screen to switch between different studies. This could possibly be visualized by using large thumbnails of preselected (pre-operative) studies. Also an extra screen could be added in which radiologists can access miscellaneous settings such comparing with other patient information. Although the results are generally positive in favour of performance and experience evaluation in the radiology room, results need to be interpreted with caution, as multimodality of medical imaging for the breast cancer diagnosis interaction could also be benefiting from a novelty effect. Further experiments with different systems and possibly during real breast diagnosis procedures, will have to demonstrate whether this is the case. As suggested above, it would be very interesting to test several interaction systems in a comparative performance and experience evaluation. This might yield interesting insights and possible overlooked usability issues.

6. Conclusions

There is a lot of information concerning UIs for clinical purpose. However, the effort for the development of a UI in a breast multimodality screening context is still scarce. This work fill the above gap, presenting innovations in the field. More specifically, this work presented an effective way at developing a new UI to collect large amount of annotations that characterises the lesions under inspection (i.e. masses and calcifications). Furthermore, this collection is performed by several radiologists that can interact simultaneously/remotely with the developed UI. This is a significant step towards an efficient way to collect data, since it is possible to collect lesion information in a very short time. Nevertheless, the diagnostic process in medical imaging is mandatory for the correct detection of breast lesions, inspection by visualization and annotation of lesions in multimodality view is therefore fundamental for the purpose.

An empirical methodology was used on this work. It has shown to be successful in detecting the needs of radiologists, indicating the appropriate testing interaction system, performance measures and finally determining usability improvements. It provides a practical methodology that can be applied to new practical medical interaction innovations, which will certainly become more and more important in the years to come. While it is a welcome addition to the generally slim amount of HCI and UCD medical UIs research in general. We also applied demographic and psychometric methods throw the analysis and evaluation of questionnaires, understanding user profile and characteris-
tics, to measure the user needs for satisfaction with system usability and UX. This brought us the information about most useful tools to our UI and what were the priorities of the features. Surveys and quantitative measures are analyzed as conventional descriptive statistics regarding this information. Meaning is added from that outcomes to the qualitative findings, providing indicators of feature positioning and interaction aspects. We worried about how prototypes could actually be useful to the intended radiologists. The prototype phase gave us the knowledge over our system and our research study. Once this work goals have been completed we would have a specification of the system architecture with expanded capabilities. Thus enabling web services using our system source with several core services and the web browser application.

Acknowledgements
To my family, friends and colleagues.

References


