Pós-Graduação em Ciência da Computação

CLEDJA KARINA ROLIM DA SILVA

DESIGN GUIDELINES FOR GENERATING AUGMENTED REALITY INSTRUCTIONS

Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

RECIFE
2016

Cledja Karina Rolim da Silva

Design Guidelines for Generating Augmented Reality Instructions

Este trabalho foi apresentado à Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutora em Ciência da Computação.

ORIENTADORA: **Prof.ª Veronica Teichrieb**

RECIFE
2016

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

Silva, Cledja Karina Rolim da
Design guidelines for generating augmented reality instructions / Cledja Karina Rolim da Silva. – 2016.
113 f.: il., fig., tab.

Orientadora: Veronica Teichrieb.
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2017.
Inclui referências.

1. Visão computacional. 2. Realidade aumentada. I. Teichrieb, Veronica (orientadora). II. Título.

006.37 CDD (23. ed.) UFPE- MEI 2017-107

Cledja Karina Rolim da Silva

Design Guidelines for Generating Augmented Reality Instructions

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Ciência da Computação.

Aprovado em: 04/03/2016.

____

Orientadora: Profa. Dra. Veronica Teichrieb

BANCA EXAMINADORA

____

Profa. Patricia Cabral de Azevedo Restelli Tedesco
Centro de Informática/ UFPE

____

Profa. Carina Frota Alves
Centro de Informática / UFPE

____

Prof. Dieter Schmalstieg
Institute for Computer Graphics and Vision/ Graz University of Technology

____

Prof. Joao Marcelo Xavier Natario Teixeira
Departamento de Estatística e Informática / UFRPE

____

Prof. Walter Franklin Marques Correia
Departamento de Design / UFPE

ACKNOWLEDGEMENTS

I would like to thank you my supervisor Veronica Teichrieb for comments, critics, support and direction during all this time.

Thanks my colleagues and friends from Voxar Labs, your help and enthusiasm made this work a little bit less hard.

During my PhD work I had the great opportunity to perform an internship in the Institute for Computer Graphics and Vision at Graz University of Technology. I am grateful for this enriching opportunity that Professor Dieter Schmalstieg gave me by accepting to be my co-advisor and by allowing me in his group. There I was very well received by his group, and I would like to thank all, especially Denis Kalkofen for all support provided.

The internship was possible due to the CNPq scholarship through the program “Ciência sem fronteiras”. This PhD work was also supported by the CNPq PhD scholarship. I am grateful to this institution for the financial help.

I would like to thank the Alagoas Federal Institute for the support.

Finally, but not less important, I would like to thank my family, which are always on my side motivating and helping me.

ABSTRACT

Most work about instructions in Augmented Reality do not follow established patterns or design rules – each approach defines its own method for conveying instructions. This work describes our results and experiences towards defining design guidelines for Augmented Reality instructions. From these guidelines, we propose a set of instructions and empirically validate them. The guidelines were derived from a survey of the most common visualization techniques and instruction types applied in Augmented and Mixed Reality. Moreover, we studied how instructions were done in 2D and 3D and how they can be applied in the Augmented Reality context. We observed that most work is related to object instructions and less work to body movement instructions. However, our proposal can be used in both cases. To validate our proposal we implemented a C++ system, which -can be seen as a library to be used in different kinds of environments where the instructions of body and objects movements are important. A RGB-D sensor was applied to capture the movements. As result, we had visualization techniques applied together with typical AR instructions and indications of what kind of instruction could be used to: emphasize parts, indication of direction of the movement, management of occlusion, management of depth and feedback.

Keywords: Visualization. Instruction. Mixed and Augmented Reality. Design Principles.

RESUMO

Grande parte de trabalhos relacionados a instruções em Realidade Aumentada não segue padrões ou guias de desenvolvimento – cada abordagem define seu modo próprio de transmitir instruções. Este trabalho descreve nossos resultados e experiência na direção de definir guias de desenvolvimento ou projeto para aplicações que utilizem Realidade Aumentada. A partir dessas guias de desenvolvimento, propomos um conjunto de instruções e empiricamente validamos as mesmas. Essas guias foram derivadas de uma pesquisa extensiva sobre as técnicas de visualização e instruções relacionadas à Realidade Aumentada e Mista. Além disso, verificamos trabalhos sobre instruções 2D e 3D para entender o funcionamento desta área para checar como e que tipo de informação poderia ser adaptada e usada no contexto de aplicações para Realidade Aumentada. Observamos que a maioria dos trabalhos é relacionada a instruções com objetos e há poucos trabalhos relacionados a instruções de movimentos do corpo. Para validar nossa proposta, implementamos um sistema em C++, o qual tem o objetivo de ser uma biblioteca para ser usada em diferentes tipos de ambientes ou contextos onde instruções de movimento do corpo sejam importantes. Para capturar o movimento do corpo, um sensor RGB-D foi utilizado. Como resultado, apresentamos técnicas de visualização aplicadas com instruções comuns em Realidade Aumentada e indicação de que tipo de instrução pode ser usada para: realçar partes, indicação da direção do movimento, gerenciamento de oclusão, gerenciamento de profundidade e retorno para o usuário.

Palavras-chave: Visualização. Instrução. Realidade Mista e Aumentada. Princípios de Desenvolvimento.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Over the last years the Augmented/Mixed Reality area has received a lot of attention due to the range of applications that may benefit from its use, the increase of processing power and the price reduction of the hardware. In this work, we will use the concepts presented by MILGRAM and KISHINO [1994] for defining Augmented Reality (AR) and Mixed Reality (MR); those concepts are illustrated in Fig. 1. AR is seen as part of the MR area and it deals with the process of augmentation or replacement of some aspects from the reality; Augmented Virtuality deals with the merging of real objects into virtual environments. In the extremes of the taxonomy schema, there are the real and virtual environments. MR is a general area that deals with the process of augmenting, removing, and changing real and virtual objects in real and virtual scenes. For this work, we will focus on visualization issues related to augmenting the scene, that is to say, AR approaches.

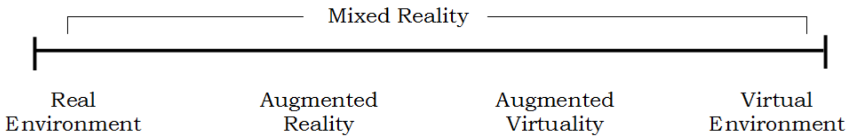

Fig. 1. Taxonomy from MILGRAM and KISHINO [1994]

The development of AR applications requires taking different aspects into consideration, such as depth perception, occlusion between virtual and real objects, environment tracking, user interaction, and so on [KRUIJFF et al., 2010]. One major issue is how to present information to the users correctly, so as to help them during their tasks. This entails considering context and users goals when presenting information.

Most of the visualization approaches applied in AR are adapted or improved from the traditional 2D/3D Information Visualization area [KALKOFEN et al., 2011]. However, it is not simple to apply them, since there are specific challenges to overcome such as, real-time runtime, dealing with virtual and real contexts at the same time, and considering the user context and goals. Besides, many AR applications have been developed to fit a specific scenario/context/task the original one. For example, an AR application that is developed to instruct a user to assembly a table has particular features (as the way of getting the data and the way to present the instruction) that can be difficult to apply in assembling building blocks. Usually, a lot of changes must be done to adapt the applications. And there are others that are virtually impossible to run in a different scenario.

In this thesis, an *instruction* is a way to teach, guide or orient users during a task execution – the task must be clear and self-explanatory, in a way that the user does not have doubts about how to do it. It is atomic. Typical examples of AR instructions are: text, animation, 2D/3D virtual objects and images. As there are different kinds of instructions (described further in the next chapters), a set of instructions can be used together according to the task to be done. SELIGMANN and FEINER [1991] argue that an illustration (images, draws, images) can fulfil a communicative intent such as showing the location of an object or showing how an object is manipulated.

In this thesis, the task is related to a training process[1] and it can be associated to objects and/or body. Examples of training with objects are maintenance and installation; examples of training with body are: body workouts, physiotherapy movements and dance.

The process of creating instructions to be followed by others has been done for a long time, from the most basic ones, such as printed manuals, to videos or 3D instructions [ADDISON and THIMBLEBY, 1995; YAMADA et al., 1997]. The use of AR for instructions can improve significantly the performance of the users, because differently from the traditional approaches where the user changes his attention between the task and the instructions, with AR, the user can focus his attention on the task and the manipulated objects simultaneously with the instructions provided on top of the object of interest.

DUH and BILLINGHURST [2008] made a survey from 1998 to 2007 about the main subjects cited in the ISMAR conference (International Symposium on Mixed and Augmented Reality), and presented the amount of visualization related papers. There are not many, but the number has been increasing during the years - this is illustrated by the red line in Fig. 2. We intend to analyze the visualization techniques presented and how they can be applied to instruction procedures.

---

[1] It is possible to call learning process, but we consider the term 'training' more adequate in this thesis

Year	98	99	00	01	02	03	04	05	06	07		
Category											To-tal	%
Tracking	6	6	2	7	7	5	9	5	8	9	63	20.1
Interaction	2	9	2	6	3	1	3	8	9	7	46	14.7
Calibration	5	6	4	5	6	3	2	1	3	6	44	14.1
AR App.	6	7	2	9	5	8	2	2	1	4	45	14.4
Display	0	4	5	7	2	3	3	4	1	8	37	11.8
Evalua-tions	0	4	1	3	2	2	0	3	5	4	18	5.8
Mobile AR	1	0	1	1	0	1	1	1	3	4	19	6.1
Authoring	0	0	0	1	2	3	3	2	0	1	12	3.8
Visualiza-tion	0	0	0	2	1	3	0	2	3	5	15	4.8
Multimodal AR	0	2	0	0	0	0	1	0	3	2	8	2.6
Rendering	0	2	1	2	0	1	0	0	0	0	6	1.9
Total	20	40	18	43	28	30	24	28	35	47	313	100

Fig. 2. Main subjects cited in ISMAR from 1998 to 2007

1.1 PROBLEM STATEMENT

Usually each major issue in AR (for example, scene distortion or depth perception) is tackled using a specific visualization solution. We will argue in the next chapters that there is no agreement in the literature about the best way to present information and mainly, for systems that deal with education and training issues, there are no rules and guidelines about how the information must be presented to users to help and conduct them during the use of the AR system.

To give an instruction, AR developers need to decide what kind of visualization technique can be applied. As there are many visualization techniques that can be applied and changes in the real and virtual contents that can occur, it is difficult to choose and apply them in the user instruction process (how to choose the techniques that can support the communication of the instruction/information). This process involves the study of different subjects like usability, aesthetics, context, and perception issues. Often, traditional 2D/3D visualization techniques are adapted or improved to be applied in AR/MR area, so when these subjects area processed rightly can benefit others areas. But, some of these subjects are also still unsolved in 2D/3D visualization areas, as it is presented in CHEN [2005]. Although in the AR/MR area there are no similar papers, there is an increase of papers associated to visualization issues.

To apply visualization techniques to help users in their tasks involves a lot of particular issues related to each environment/context. It is easy to verify that most

AR applications are developed for specific scenarios. For example, ANDERSEN et al. [2009] give an AR instruction system to fix a pump; the approach by REINERS et al. [1998] presents a complete system to fix a door lock.

Furthermore, the choice of AR developers is not limited by the kind of visualization technique, but also by the many possible kinds of instructions that can be applied within the scope of AR approaches.

1.2 Research Question

The research question of this thesis is: *is it possible to define guidelines related to the process of user instruction in AR?*

Together with the answer, we would like to help the AR developers in the process of creating AR approaches related to instruction. Different from the usual design patterns that have solutions to common code problems in object-oriented project development [LARMAN, 2004], the solution proposed is not related to code, but an orientation to guide them on issues that improves the use and the application of instructions by common users.

1.3 Objective

In an attempt to answer the research question stated above, perceptual issues must be considered to give the correct instructions and to define the best suited visualization technique. We will present them in the next chapter. This thesis has the goal to study patterns in the visualization and instructions approaches in AR to create guidelines to improve the user experience in AR; to reach this goal, we have followed these steps:

- Analysis and implementation of visualization techniques to render AR scenes;
- Analysis and implementation of instructions in AR;
- Definition of guidelines for instructions in AR - focus on object and body instructions; and
- Validation of the guidelines.

1.4 Purpose of the Research

Usually, each AR application has its own way to manage instruction. AR developers have no clear guidelines to follow on the issue. Various kinds of instructions can be applied, but with no specific reason. Some common instructions used are: color changes, 3D animations, images and 3D virtual versions of a real object. Besides, AR applications can focus on body or object instructions, and the nature of both can bring doubts about whether the same kind of instructions can be applied in both cases. Thus, we have identified the necessity to define guidelines to help in the development and use of instructions in AR applications.

1.5 Thesis Outline

This thesis is organized as follows.

In Chapter 2, the concept and theory related to Functional Realism, the perceptual issues related to AR scenes and a three-stage approach that was applied to create and validate the guidelines proposed in this thesis are presented. The next chapters are based on this approach.

Chapter 3, the Identification Stage, discusses how the guidelines are identified – based on the most representative AR approaches related to visualization and instruction.

Chapter 4, Instantiation Stage, presents the guidelines and how common visualization techniques can be applied in AR instructions. Besides, an AR instruction system that applies the guidelines is presented. There are results related to body and objects instructions.

Chapter 5, Evaluation Stage, presents a pilot study applied to a group of users to validate our proposal.

Chapter 6 presents our final considerations, as well as, contributions and suggestions for future work.

2 BACKGROUND

This chapter describes the related works and theoretical foundations that are important to the correct understanding of this work. Section 2.1 describes functional realism, a kind of computer graphics variety that aims at presenting certain information crucial to the execution of a specific task to the user – and is also related to the process of giving instruction. Section 2.2 presents the underlying concepts and kinds of perceptual issues in AR. Section 2.3 presents the works related to guidelines in AR. Section 2.4 presents the methodology followed in this thesis to define and validate our guidelines.

2.1 FUNCTIONAL REALISM

One of the goals of this thesis is to study visual issues that can affect the perception of AR scenes, to improve the user experience when performing a task. Considering visual tasks, it is really important that the user is able to make reliable visual judgments when observing a scene. This happens when the information provided by the scene and its augmentation represent the knowledge about the meaningful properties of objects in the scene, such as their shapes, sizes, positions, motions and materials.

Functional realism is one of the three varieties of realism in computer graphics identified by FERWERDA [2003]; it is concerned with enabling the user to perform a real world task appropriately. It that the image generation process produces an image (the augmented content in an AR context) that provides the same visual information as the scene, allowing a user to perform useful visual tasks.

Typical factors considered in the execution of a task are its goal and context. The context describes the conditions or state of the environment and the objects needed? to realize the task. Both goal and context can change the environment and objects (real and virtual) issues (appearance, state, location). It might be the case that each factor is dependent and influences the other.

There are works that try to give some directions about tasks and AR/MR. Examples of those are the approaches from TORY and MOLLER [2004] and ELMQVIST and TSIGAS [2008]. The first argues that the effectiveness of the visualization depends on perception, cognition, and the users' specific tasks and goals. The second proposes a taxonomy for 3D occlusion visualization and presents three visual perception tasks where occlusion can happen: target discovery, target access, spatial relation. If there is full occlusion of a target the task will be hampered. Partial occlusion can have different effects: the task cannot be done or it can be done with difficulty. Their taxonomy considers seven properties to the design space for the occlusion management techniques: primary purpose, disambiguation strength, depth cues, view paradigm, interaction model, target invariance and solution space.

The other two varieties of realism in computer graphics are physical realism and photo realism. In physical realism the image provides the same visual stimulation as the scene and in photo-realism the image produces the same visual response as the scene.

WARD [1989] illustrates an example of physical realism; his work describes a physically-based rendering system tailored to the demands of lighting design and architecture – an example from his approach is presented in Fig. 3, a visualization of the illuminance levels of a room. Since physical realism is very hard to achieve in real-time, this variety is not approached in this thesis.

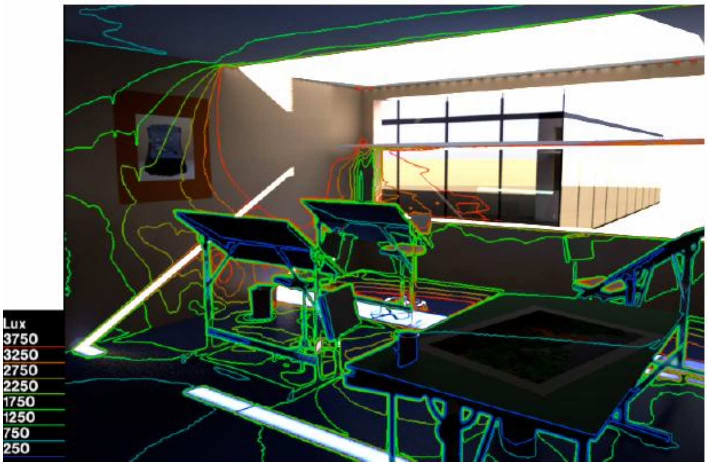

Fig. 3. An example of physical realism: a visualization of the illuminance levels of a room from WARD [1989]

Fig. 4 illustrates an example of photo-realism obtained from SANTOS et al. [2012]. The black chess pieces and the teapot are virtual objects, and the others are real. Since the visual response of the image is very close to the one from the scene it is difficult for ordinary user to realize (sometimes even experts) that the virtual objects are not real. On the other hand, Fig. 5 presents a virtual representation of a car piece on top of the real car to guide the user to remove it during a specific maintenance procedure. This example is a project from BMW SERVICE [2014]. In this case, the virtual object does not look like the real one, but its properties (shape, size, position) are close enough to the real one to ensure that the user understands that it is the correct piece to remove from the car.

Fig. 4. Example of photo-realism: virtual and real objects are placed consistently; users have difficulty to realize the difference between the real and virtual ones

Fig. 5. Example of functional realism: virtual objects (hands and car piece) are used to help the users during a task

This thesis will investigate visualization techniques for instructions in AR that focus on both photo and functional realism.

2.2 PERCEPTUAL ISSUES IN AUGMENTED REALITY

KRUIJFF et al. [2010] proposed a classification of perception issues in agreement with the areas related to the process of augmenting a scene, called perceptual pipeline. They classified perceptual problems in three great areas: scene distortions and abstraction, depth distortions and object ordering and visibility.

- Scene distortions and abstraction deals with problems related to object recognition, size perception and visual segmentation. Fig. 6, obtained from KALKOFEN et al. [2009], presents an example: the virtual engine is presented in red color; if a real car is shown with the same color (on the left), the user has difficulty to perceive the car engine. But, when the real car has a different color (the yellow car on the right image), the user has less difficulty to perceive the relationship between real and virtual objects;
- Depth distortions and object ordering: it deals with problems related to depth ordering, such as occlusion and overlaid information. An example of these issues is presented in Fig. 7 from PADILHA et al. [2013]: on the left image, the virtual pipes (in green color) are only overlapping on the scene; on the right image, the occlusion is managed;
- Visibility: deals with screen problems, such as color, texture and brightness. A way to deal with visibility problems, such as texture or features of real objects, is the use of Diminished Reality [ROLIM and TEICHRIEB, 2012]. In this sense, an modified virtual version of the object is placed to improve the user experience, as we can see in Fig. 8, adapted from LEAO et al. [2011].

Fig. 6. On the left image an example of scene distortion problem; on the right image a possible solution

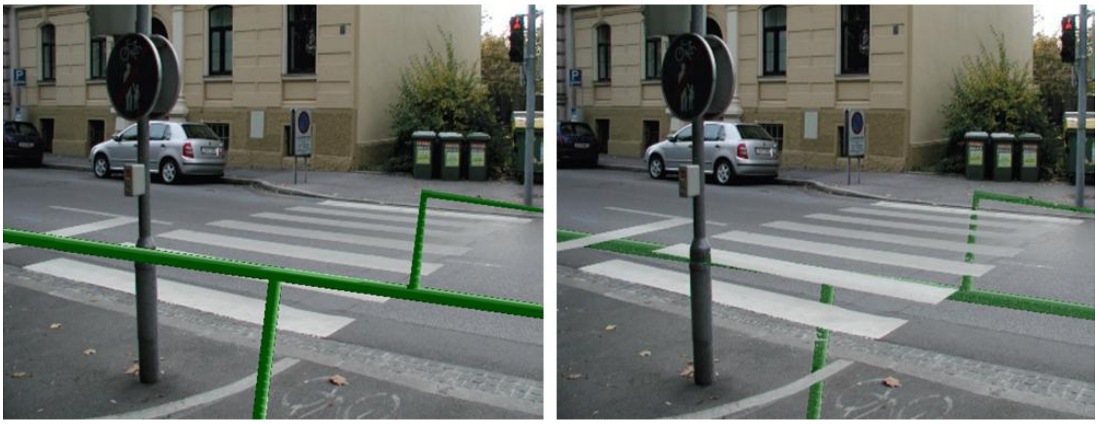

Fig. 7. An example of management of occlusion by the approach from PADILHA et al. [2013]

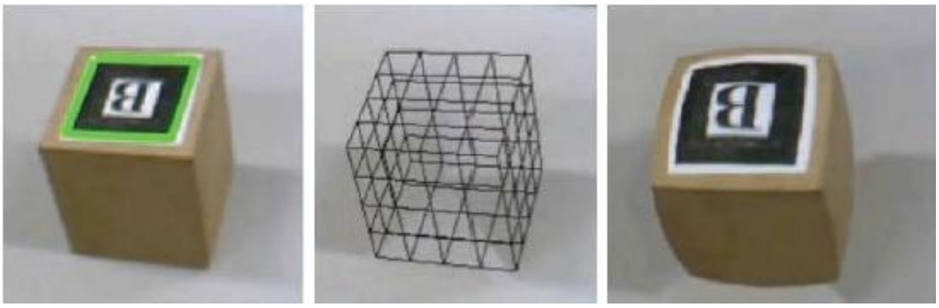

Fig. 8. An example to solve texture and undesired features/objects on real scenes: the application of Diminished Reality, a real object (left image) is removed to put in its place, an improved version of it (right image)

As this thesis works on visualization and instructions in AR, these problems must be managed to give a better answer to users. So, we try to manage and solve some problems related to this perceptual pipeline in the instruction process.

2.3 RELATED WORK

We have searched for guidelines or principles for presenting AR-based instructions and visualization techniques applied to instructions. We found few works in this line which are discussed next.

The work from NEUMANN and MAJOROS [1998] is an attempt to define rules for creating instructions in an augmented scene. The authors present cognitive issues in the design of virtual content for manufacturing and maintenance tasks. They define that we must have specific designs or objects to focus user attention on (warnings or cautions), design objects to be adjustable, make objects dependent of operating conditions (e.g., higher contrast callouts in bright viewing conditions), so that the users can have the possibility to 'copy' and 'paste'. Their rules are specific to objects and in some cases are difficult to apply them in other scenarios (for example, body instructions). Our goal is to define more general rules (the guidelines), but we have similar point: the guidelines must have a way to focus the user attention to part of the scene/object/body to help the user to understand which part is necessary to move/change/replace.

FURMANSKI et al. [2002] define a set of guidelines to manage the depth perception and visual complexity to x-ray vision in AR. X-ray is a typical technique used in 3D visualization that was adapted to be applied to AR applications. It is an interesting technique to manage occlusion. In spite of presenting interesting points and guidelines (for example: distance conveyance, set up correctly the real scene that the information will be presented in), they are too specific with the application of the x-ray technique. In our proposal, we apply the x-ray technique as an instance of a guideline to manage occlusion.

KALKOFEN et al. [2011] present a set of visualization techniques that deal with depth perception cues, such as occlusion and relative size, to improve the comprehension of the relationship between virtual and real contents. Between the visualization techniques applied there are edges emphasis and x-ray. We applied both visualization techniques as a way to instantiate our guidelines.

GILMENO et al. [2013] propose a framework to develop AR applications to object instructions in industrial procedures. This framework allows non-programming users to develop assembly and maintenance tasks. It is an interesting framework, with a pre-defined set of instructions. Comparing the framework with our proposal, we have a pre-defined set of instructions, but defined in agreement with our guidelines. Besides, our guidelines are more general, and they can be applied in other contexts.

ANDERSON et al. [2013] define guidelines for movement training systems based on their experiences and review of the literature; the guidelines are: leverage domain knowledge, motivate the user, simple presentation, low cognitive load, adaptive guidance, summary feedback, user-driven learning. Their guidelines do not focus on the way to give user instructions; our proposal focuses on movement guiding and specific issues related to the presentation of information in AR scene.

TANG et al. [2015] present an AR approach to teach physiotherapy movements at home. To track correct the body movements, the users need to wear specific body sensors. They present interesting kinds of instructions, such as arrows, angles and multi-view (there are more than one RGB camera to visualize the movement). To check the user movement an error metric was implemented. The way that they present the instruction is simple and easy to users understand the movement range, but improvements need to be made with respect to depth and occlusion perception. We have applied some of their instructions (arrows, colors and text) in our proposal as a way to validate our guidelines.

There are many contributions from other areas to AR (for example computer vision), especially many adaptation of visualization techniques to AR. We surveyed approaches from this area, to verify interesting contributions to our work. One of them was the approach from AGRAWALA et al. [2011]. The authors defined a set of guidelines to create automatic instructions from the analysis of the best handmade drawings. We followed a three-stage approach defined by them in our proposal to define our guidelines (details will be presented in section 2.4).

SHANMUGAM [2015] defines five principles in the development of virtual reality applications: everything should be reactive, motions should be restricted to interaction, text and image legibility, ergonomics and sound effects. since our goal is related to how instructions could be given, the principle 'everything should be reactive' is related to our guideline 'feedback'; the principle 'text and image legibility' is related to the process of given an instruction: they must clear and easy to read, especially in AR applications where the real environment conditions (sun light or dark environment), in some cases can difficult the identification of the virtual data.Their previous work is related to general guidelines to virtual reality application, the approach from KIM and SONG [1997] presents five instructional designs guidelines specific for virtual reality in classroom applications. The guidelines are: provide divergent learning outcomes (apply real situations in virtual classroom), focus on learner-centered control (the user can customize the environment to fit personal needs and interests), provide a high level of user interaction (the application should permit and require the users to maintain an active participation), follow the principles of instructional design (refine the learning environment considering internal and external learning factors which affects successful learning) and consider constructivist learning principles (the virtual learning environment should facilitate the user's individual construction of meaning from presented stimuli). These guidelines are very appropriate to be applied in AR applications, but some of them could be hard to be fully explored by the developers, because of the lack of full control of the real environment. The main issue of these guidelines is the possibility to the application/environment to react in agreement with the user actions and personal options. We have the goal that our guidelines also try to react, and we defined a guideline that try to manage this issue (feedback).

An interesting approach to give body instructions is the 3D approach of BOUVIER-ZAPPA et al. [2007], where different motion cues are given to illustrate the correct body movement of an avatar. The motion cues are: arrows, noise waves and stroboscopic motion. Fig. 9 presents an example from their work that shows the steps of a soccer player; the arrows indicate direction and position; the color gradient indicates the velocity. As a 3D technique, they already have all the body data, and this enables the application to have full control of the data and presentation of the instruction. We applied arrow as instruction in our work, as they did, but their arrow has more functionalities, such as the indication of velocity. They used an avatar to give an overview of the movement. But, different from AR approaches, they do not manage feedback issues nor real time data.

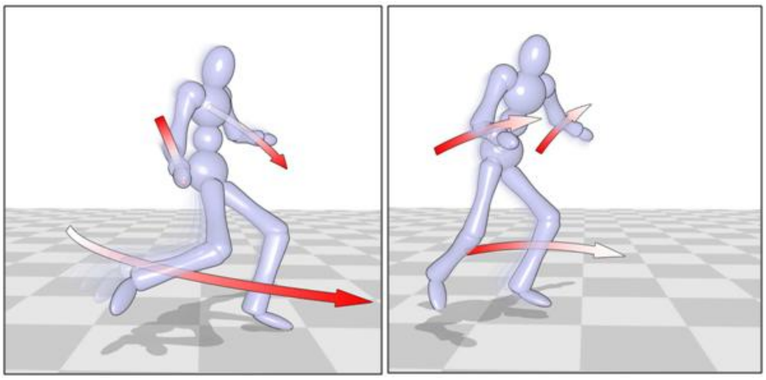

Fig. 9. An example of body instructions by BOUVIER-ZAPPA et al. [2007]

Table 1 summarizes the approaches presented in this section; the third column presents our proposal in comparison with the related works.

Table 1. Summary of the related works

Approach	Description	Our proposal
KIM and SONG [1997]	Designs guidelines virtual reality in classroom applications	Design guidelines to AR
NEUMANN and MAJOROS [1998]	Cognitive issues in design of virtual content for manufacturing and maintenance tasks	Focus on body and object instructions
FURMANSKI et al. [2002]	Guidelines to manage the depth perception and visual complexity to x-ray vision	Different kinds of visualization techniques
AGRAWALA et al. [2011]	Guidelines for creating 3D automatic instructions	Guidelines for AR instructions
BOUVIER-ZAPPA et al. [2007]	Body instructions	Body instruction to AR
KALKOFEN et al. [2011]	Focus on visualization techniques	Guidelines for instructions
GIMENO et al. [2013]	Object instructions to industrial procedures	Body and object instructions
ANDERSON et. al [2013]	Body instructions / General guidelines to movement training	Focus on movement guiding using AR-based instructions
TANG et al. [2015]	Body instructions	Guidelines for instructions
SHANMUGAM [2015]	Design principles to virtual reality	Design guidelines to AR

2.4 THE THREE STAGE-APPROACH TO CREATE AND VALIDATE THE GUIDELINES

AGRAWALA et al. [2011] define a three stage-approach for creating effective and automatic visualizations: identification, instantiation and evaluation. They are done in sequence and must be validated by the user in the end.

The authors define design principles as *prescriptive rules describing how visual techniques affect the perception and cognition of the information in a display*.

Despite the fact that their approach and examples are not specific to the AR area, but applied in 2D and 3D visualizations, this thesis will follow the same steps, due to their independence of context, enabling a developer/designer to apply them to different tasks and environments (one of our goals in this research). In our approach, we call them *design guidelines*, because we believe that it is more adequate for the development process. The stages are presented next:

1) Identification: they identified the design principles from the analysis of the features from the best hand drawings examples. In the context of assembly instructions they identified three principles: use a step-by-step sequence of diagrams showing one primary action in each diagram, use guidelines and arrows to depict the actions required to fit parts together, and ensure that the parts added in each step are visible;

2) Instantiation: to validate their approach, they apply the design principles to create cutaways [SIGG et al., 2012], exploded views [BRUCKNER and GR, 2006] and how-things-work illustrations [MITRA et al., 2010] in two different domains, cartography and technical illustration.

3) Evaluation: This stage deals with feedback from the users and how useful the instruction was. Possible evaluation methods include subjective user feedback (interviews and surveys) and user studies comparing the performance of users with the the AR system those with paper-based instructions.

Fig. 10 illustrates the AGRAWALA et al. [2011] process to define their guidelines based on the best handmade illustrations.

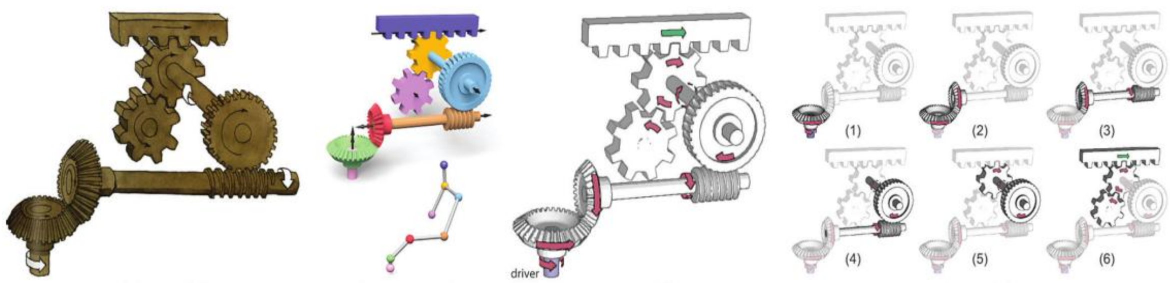

Fig. 10. Approach from AGRAWALA et al. [2011]: from a handmade design (first draw) automatic illustrations are created: motion analysis (second image), annotated illustration (third image), and parts emphasis and movement sequence (fourth image)

2.4.1 *METHODOLOGY*

To create our guidelines, we followed the same steps proposed by Agrawalla et al. [2011] but with different content and processes in each stage (their approach focuses on 3D visualizations, while ours focuses on the AR area), so this is another contribution of this thesis.

The pipeline and stages defined by the authors are illustrated in Fig. 11: the first stage is the identification, second is the instantiation and finally, the evaluation. The next three sections give an overview of the content and steps defined in each stage to facilitate the understanding of the next chapters (in which we explain each step in further detail).

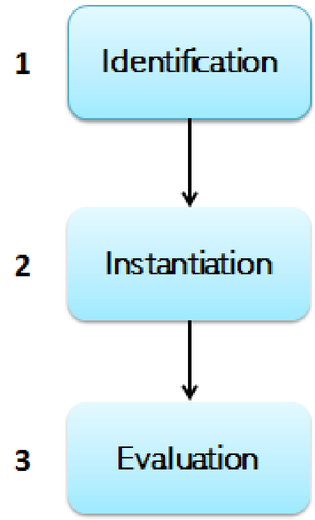

Fig. 11. Three-stage approach followed in this thesis

2.4.1.1 IDENTIFICATION

Agrawalla et al. [2011] defined their guidelines based on the analysis of the best handmade drawings. In this thesis, the guidelines were defined from the analysis of visualization techniques (step 1) and instructions techniques (step 2). These techniques were analyzed with the assumption that they were meant to give instruction or information to users. These two steps are illustrated in Fig. 12. In each step, the AR approaches are organized into categories to facilitate the identification and application of them as part of the instruction process. In the visualization techniques analysis three categories were identified: visual photometric consistency, visual geometric consistency and visual attention. In the instruction techniques analysis, two categories were identified: classical mode and perceptual mode. This stage, its categories and the guidelines proposed are presented in details in Chapter 3.

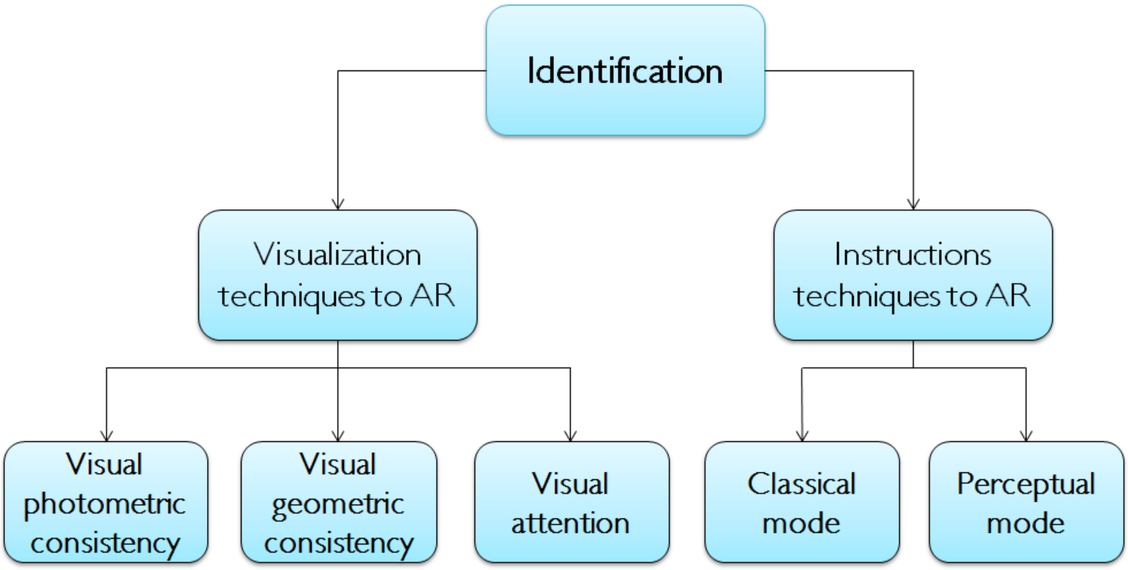

Fig. 12. Steps from the instantiation stage

2.4.1.2 *Instantiation*

In this stage, prototypes were implemented to validate the guidelines, in accordance with the obtained data from previous stage. Examples of typical visualization techniques that can be applied to give instruction or help the user during a task were also proposed. Fig. 13 illustrates in the blue rectangle the category identified in this work and in the gray rectangles a proposed visualization technique that can be applied to give instruction.

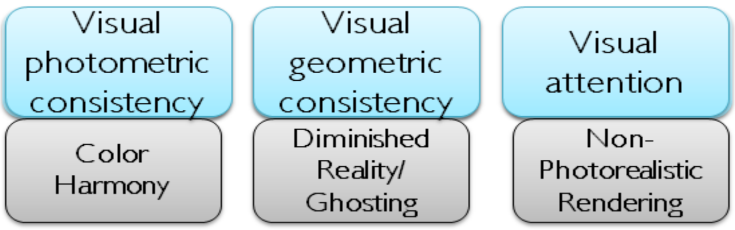

Fig. 13. Examples of typical visualization techniques that can be applied to the user instruction

An instruction for each guideline proposed was implemented. A summary of those is presented in Fig. 14 (blue rectangle a guideline, in gray rectangle an example of instruction that can be applied). The prototypes implemented and application of visualization techniques as instruction is presented in Chapter 4.

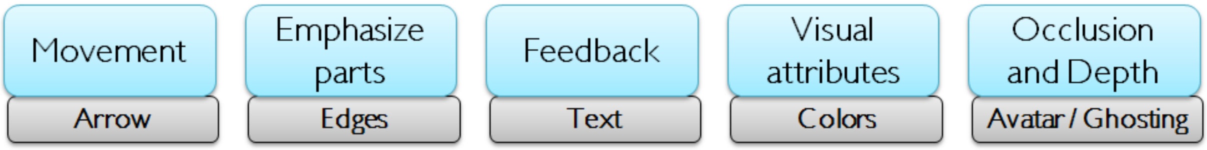

Fig. 14. Examples of possible instructions to be applied in each guideline proposed

2.4.1.3 Evaluation

To validate the guidelines, three steps were defined in this stage: first one an analysis of answers of instructors was carried out, second, a pilot test was applied to users, to check if the guidelines were useful in a defined task; and third, we analyzed answers from AR designers/developers related to our guidelines in their AR applications. These steps are presented in Fig. 15. This stage is presented in details in Chapter 5.

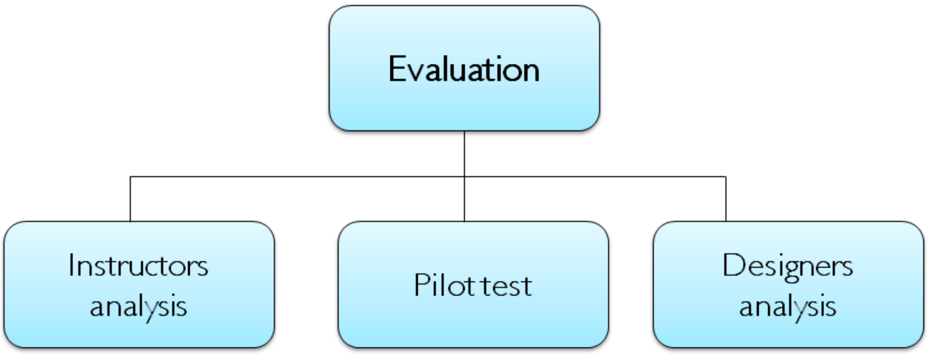

Fig. 15. Steps of evaluation stage

Each stage presented will be explored in details in the next chapters.

3 IDENTIFICATION STAGE

The identification stage in our approach was defined by the analysis of the techniques related to visualization and instructions in AR. Section 3.1 presents our survey about visualization techniques applied in AR as a way to understand how they are used and applied to present instructions; Section 3.2 presents the study regarding instructions in AR; we analyzed the types of instructions applied, if they are related with body or object instructions and if they manage some perceptual issue. In the end of each section, we present our conclusions and observations that were used to define our guidelines in Section 3.3.

3.1 VISUALIZATION TECHNIQUES IN AUGMENTED REALITY APPLICATIONS

Many visualization methods and techniques applied to AR applications are inherited from 2D and 3D visualization approaches. Visualization issues are one of the most important subjects in AR applications because of their significant impact on the usability factor of the systems, mainly the ones dealing with education and training [PUIG et al., 2012; FARINAZZO et al., 2013]. Besides, to apply AR approaches in different contexts is a challenging problem to tackle. This chapter presents a survey of the main techniques applied in AR and MR that can be applied to AR instructions.

In our research, we searched for representative techniques for each one of the three perceptual issues category presented in previous chapter. In spite of having a many AR/MR approaches to help users during a task or a procedure, this chapter tries to identify how visualization techniques are applied as part of the process of user instruction.

In this work visualization techniques are classified into three categories: *visual attention*, *visual photometric consistency,* and *visual geometric consistency*. The visual attention category deals with techniques that aim to focus the user attention in parts or objects (real or virtual) on the scene. The visual photometric consistency brings out the techniques that deal with photometric issues. Finally, the visual geometric consistency category presents the visualization techniques that deal with the correct placement of virtual objects and the removal of real objects. The following sections present each category in further detail.

3.1.1 VISUAL ATTENTION

In this category, there are techniques that try to capture the user attention to parts of the scene to guide the user. Some of them applied the theory or computer models from the Computer Vision Area [SZELISKI, 2010], that simulate and/or understand the Human Visual System (HVS) to identify relevant and important areas of the scene [BORJI and ITTI, 2013; FRINTROP et al. 2010].

The understanding of the HVS by psychologists and neurobiologists is used as a basis for building computational models that catch users' visual attention [BORJI and ITTI, 2013]. Many experiments have demonstrated that only a small region of the scene is analyzed in detail at each moment: the region that is currently attended. Despite the fact that users can realize the entire scene, they only focus on specific regions [FRINTROP et al. 2010] – these specific regions are defined by, for example, shape, color, tone that are different or contrast with other regions of the scene (one or more), and by the goal (for example, if the user wants a red ball, his or her attention will focus on red, round shaped objects). This selection mechanism is defined as *visual attention* and the specific regions are called *salient*. The main focus area is called *overt attention* and the peripheral area seen also by the user eyes is called *covert attention*. There are two basic models of attention: *bottom-up* and *top-down*.

The bottom-up model is driven by the characteristics of a visual scene, like color, contrast, intensity [BORJI and ITTI, 2013]. These characteristics pop up in the eyes - a red ball in a white room has a huge impact in the user attention, so the bottom-up model is automatic and reflexive. The top-down model is goal oriented and uses cognitive factors, such as previous knowledge about the scene or task (for example, the user has the goal to find a key in a room; what kind of factors will he consider? Will it be the same factors for all users?).

The bottom-up model has been more studied than the top-down one; a reason for this is that the data to analyze are coming from the visual scene and so are easier to control than cognitive factors [FRINTROP et al. 2010]. There are many studies to define the features used to visual attention; most of them agree that basic features are color, motion, orientation and size of the objects in the scene. These features are computed to produce a saliency map – representation of the areas that most attract user attention. A large study about computational models of attention can be found in BORJI and ITTI [2013]. It is easy to realize that most computational systems rely on a bottom-up model, because only local and visual information are used to produce a saliency map.

Only a few AR/MR works address visual attention. One of the main reasons for this is the processing time required by the models of attention. These models have been studied and produced in the Computer Vision area, but few can run in real-time [FRINTROP et al. 2010]. Commonly the techniques applied are changed according to the proposed models to be possible being executed in real-time or dedicate hardware are applied. The techniques presented in this section use the bottom-up model. We did not find AR/MR works related with top-down models and the approaches which applied bottom-up models are few and limited. However, we believe that advances in the study of top-down models will increase their use in AR/MR approaches.

SANDOR et al. [2010] utilized a bottom-up model to improve their x-ray visualization technique. In their previous work [AVERY et al. 2009], they utilized only edge information to define the parts of the scene to be preserved and excluded. To keep a coherent visualization, they improved the technique through a saliency map, considering hue, luminosity and motion information. They computed the saliency maps from the occluded and occluding parts of the scene and made a combination of them to define the areas to be preserved. Some necessary improvements pointed out by the authors are related to bright foregrounds and refinement of the motion-based saliency. Fig. 16, on the left, presents an example of x-ray visualization with the saliency map; on the center, the same scene visualized utilizing only edge information; and on the right, the saliency map used to produce the result on the center. It is possible to see that the most outstanding objects present in both images (left and right) are the umbrellas, and they are kept with the use of the bottom-up model of attention.

The limitation of the movement detection presented before is handled by PADILHA and TEICHRIEB [2014]. They introduced an algorithm to classify parts of the frames in agreement with the motion pattern. The algorithm creates a saliency map in agreement with the movement information.

Fig. 16. The x-ray visualization technique from SANDOR et al. [2010]

There are different approaches of bottom-up models to create a saliency map. One of them applies lightness opponency, red-green color opponency and blue-yellow color opponency [ITTI et al., 1998]. These techniques are used in the work of MENDEZ et al. [2010] to emphasize an object defined previously by the user. Their work is interesting because there is no lack of context, there are minimal changes in the original image and there is an emphasis on the important objects. Fig. 17 illustrates their work, showing that the context of the real scene (left image) is preserved and the user attention is directed to a specific window (right image).

Fig. 17. The focus and context technique from MENDEZ et al. [2010]

In the work of CHEN et al. [2011a] the saliency map is used to generate an importance map that directs the use of multiple rendering styles. The input can be images, videos, 3D scenes or AR or MR scenes. In agreement with the saliency map, importance map and the user defined parameters, different rendering styles can be employed in different parts of the image/scene to emphasize diverse aspects, as it can be seen in Fig. 18.

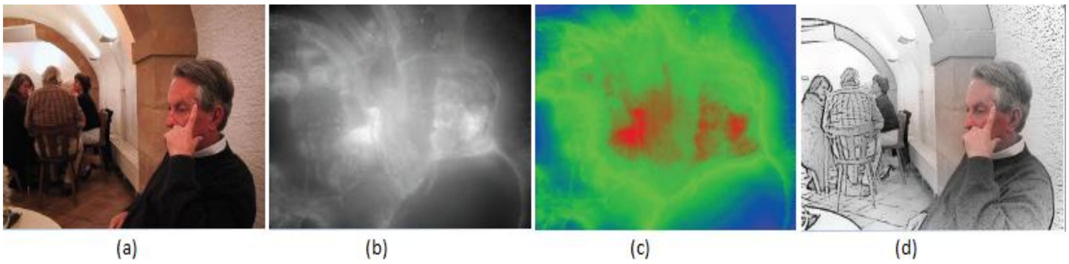

Fig. 18. The painterly rendering algorithm from CHEN et al. [2011a]: (a) the input image, (b) the importance map of the input image, (c) different weights can be given to importance map creating another importance map, (d) a rendering style is applied in agreement with the weight importance map

Non-Photorealistic Rendering (NPR) is another way to focus on important features on the scene. It is a kind of rendering style guided by artistic processes; each artistic style has the goal to emphasize certain features, communicate or simplify different ideas in agreement with the human understanding. HALLER [2004] believes that non-photorealistic pictures can be more effective at conveying information, more expressive and more beautiful, because the virtual objects don't need to be realistic, but NPR must give information. GOOCH et al. [2010] present challenges in the NPR area; one of them is the use of cognitive principles, and this is also a challenge in the MR area. KYPRIANIDIS et al. [2013] give a survey about NPR techniques for images and videos; some of them have already been used in the MR area.

NPR has been studied in the visualization area for a long time, and it has been adapted to MR. Usually, not only virtual objects are stylized, but in most cases, the entire scene is. Whenever the entire scene is stylized the process is called stylized MR and the real scene and virtual objects become less distinguishable. An example can be seen in Fig. 19 from FISCHER et al. [2008]; on the left, a virtual yellow cup is added to a real scene; on the right, the entire scene is stylized, reducing the difference in appearance between real and virtual objects.

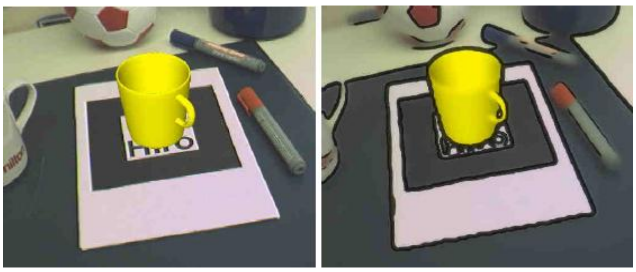

Fig. 19. An example of a stylized augmented scene: on the left, a normal augmented scene and on the right, a stylized scene by the approach from FISCHER et al. [2008]

FISCHER et al. [2008] presented a survey about techniques and applications related to stylized MR; they also presented a pipeline for this kind of work. Some works related to creating MR stylizations are CHEN et al. [2011], WANG et al. [2010], HALLER and SPERL [2004], HALLER et al. [2005] and FISCHER et al. [2005].

Silhouettes or contours are interesting resources to be utilized when creating a stylized MR application and represent a way to reduce the complexity and the identification of important areas in real scenes and virtual objects. HANSEN et al. [2010] argue that the classical rendering styles (Fig. 20, left) can occlude important information or that the use of transparency cannot convey depth information. To deal with these problems and the illustration of important and risky areas in order to support surgical decisions in liver surgery context, their work stylizes only the virtual object and it presents an advanced silhouette algorithm to convey depth and spatial context information (Fig. 20, right). Their approach is termed distance-encoding silhouettes. Besides, three visualization scenarios were presented to test the usability of the technique.

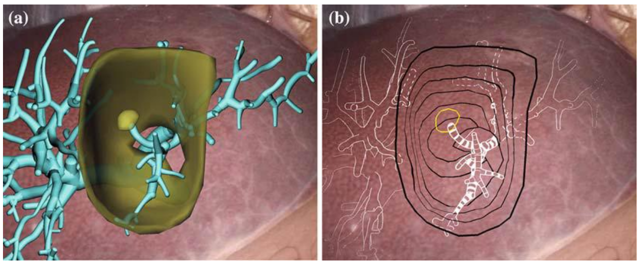

Fig. 20. NPR through silhouettes or contours: on the left, a typical augmented scene; on the right, the technique proposed by HANSEN et al. [2010]

3.1.2 VISUAL PHOTOMETRIC CONSISTENCY

This category includes the techniques that work with photometric issues, such as shadowing and chromatic adaptation. Illumination is a longstanding concern in Computer Graphics; usually it is applied in photorealistic rendering. However, it can affect the perception of any scene (photorealistic or not). In AR/MR scenes the geometry and texture of virtual objects is known, but to retrieve this information from the real environment is typically difficult. This is important in order to allow the virtual world illumination to reproduce the real sources of light and effects produced by them (e.g.: shadows, reflection, color bleeding). This category illustrates that illumination is an important issue for depth perception and relation between real and virtual objects.

MADSEN et al. [2003] and SUGANO et al. [2003] discuss the importance of shadow effects in AR. Fig. 21 (left image) illustrates the real scene, then virtual objects are added to it (center image), and the same objects are added with coherent shadows (right image). In spite of dealing with occlusion in the center image, the perception of depth, correct position and relation between the objects is enhanced in the right image due to the shadow effect.

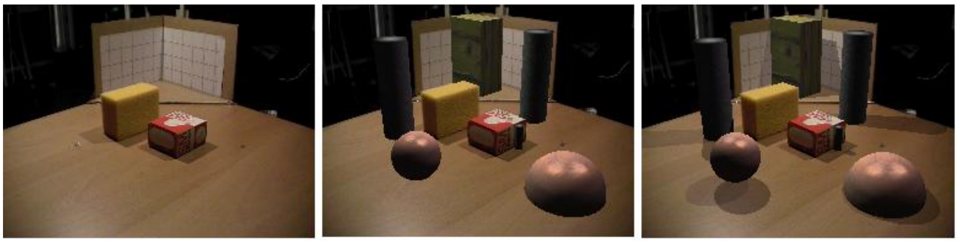

Fig. 21. Shadow effects in AR [MADSEN et al. 2003]

NOH and SUNAR [2009] give an overview about shadows techniques utilized in AR. NOWROUZEZAHRAI et al. [2011] deal with correct shadowing of virtual objects in agreement with the lightning captured from the real world. GIERLINGER et al. [2010] use different rendering techniques to deal with illumination issues in MR. JACOBS and LOSCOS [2006] give an overview and propose a classification for illumination methods. YEOH and ZHOU [2009] present a technique for realistic shadows in real-time.

The theory and use of colors have been studied in the visualization area for a long time [HUGES et al. 2013]. The perception of a scene depends on the correct choice of colors used. Human vision is affected by, for example, luminance and saturation; the size of an object can be perceived in different ways in agreement with the color used [SILVA et al. 2011]. Color harmonization is an artistic technique to adjust the colors of a given image in order to enhance their visual harmony [SAWANT and MITRA, 2008]. There are approaches related to images such as COHEN-OR et al. [2006], WONG et al. [2012] and related to videos [SAWANT and MITRA 2008]. The difficulty to use these approaches in MR scenes is the high execution time.

GRUBER et al. [2010] present a technique that automatically re-colors virtual and real world items in order to achieve more visually pleasant combination of virtual and real objects. REINHARD et al. [2004] present an approach to re-color only the virtual objects of the scene. MENK and KOCH [2011] present a system which takes into account ambient light, the material of the object’s surface, and the pose and color model of the projector to adjust the RGB values defined previously to a specific material; their work is related to spatial AR. KNECHT et al. [2011] make color adjustments in virtual objects in a way that these objects seem part of the real image. Their focus is on visualization in mobile AR, although there is the necessity to get a prior representation of the real scene, being hard to apply the technique in most MR applications. The work of LIU et al. [2009] presents a model to estimate the light information from outdoor environments, and with this information objects can be shown in a correct way. ZHU et al. [2010] utilize color information with depth obtained from stereo cameras to produce a coherent occlusion between real and virtual objects.

A good example of the importance of color information in AR scenes was presented in Fig. 6. KALKOFEN et al. [2009] present some concern about the use and application of color information in AR scenes.

3.1.3 Visual Geometric Consistency

This category includes the techniques that deal with the correct placement and identification of the virtual objects, besides the removal of real objects (or parts of the real scene) in a way to improve the augmentation. This category manages size, occlusion and texture.

As we saw in Fig. 8, sometimes it is interesting to remove real objects to put an improved virtual version of them or a more adequate virtual version of them in the environment. Besides, there are a lot of undesired objects that could cause difficulty to the user to do a task and have a complete experience with an AR system. In this sense, Diminished Reality (DR) is applied to remove real objects. In Image Processing the process of removing undesired objects is done by the Image Completion area [LIU et al. 2011], and sometimes it is called Inpaint Techniques [BERTALMIO et al. 2000]. As we have discussed before a lot of techniques usually used in 2D and 3D approaches are adapted to be applied in AR/MR areas. So, this happens in DR too; but, there are specific works developed to AR/MR, because a lot of image completion or inpaint techniques do not run in real-time and it is a very important issue in DR.

DR is seen as part of the AR research area and there are only a few related works available in the literature about this subject. It is an interesting way to change perception of real scenes, mainly because there are many possible applications, such as training (erase unimportant areas to focus on important ones improving comprehension), urban planning/architecture (erase small objects or big areas to do planning of the areas) and entertainment (situations like fights, naked people, racist comments in posters and remove undesirable advertisements in live TV sport games).

DR techniques can be multiview based approaches and a patch or fragment based ones [HERLING and BROLL, 2010]. The first approach uses different cameras to reconstruct the target area (JARUSIRISAWAD et al. [2010], SEO et al. [2008], ZOKAI et al. [2006]). The second approach uses image processing techniques to reconstruct the target area. As a main example, there is the work of HERLING and BROLL [2010], with no necessity to do any pre-processing steps and running in real-time. The process of object removal is done in two steps: the first is selection and tracking of the target and the second is image completion. Selection and tracking are performed in order to determine the part to be removed from the current image, resulting in an image called importance mask that indicates the areas to be removed and those to be preserved in the image. The object selection is done using an Active Contour algorithm. The image completion is patch-based and utilizes a bidirectional similarity measure that, in this case, verifies the level of similarity of two images. An example of result of their approach can be seen in the middle image in Fig. 22. The technique makes data and sample reductions to get real-time performance and, in some cases, their visual results are not as good as other static image approaches.

So, the authors proposed a new approach called PixMix [HERLING and BROLL, 2012]. In order to fill the target area, PixMix utilizes a combined pixel-based approach rather than a patch-based approach. This filling is defined by a cost function which is a combination of spatial cost and appearance cost (each one can have different weights). The appearance cost makes their approach become different from the similar pixel- or patch-based approaches. Other contribution of the technique is a homography-based approach that supports rotational camera movements in addition to linear camera movements. The left image on Fig. 22 presents the object selection (mask), the middle image presents the result of removing the selected object from their previous work and the right image presents the removal procedure result from the PixMix approach.

An adaption of the HERLING and BROLL [2010] approach was done by ROLIM and TEICHRIEB [2012]; they applied other kinds of image processing techniques (such as inpaint and filters), but they did not get real-time and similar visual results.

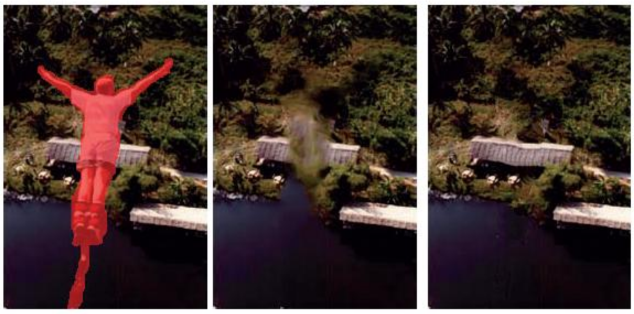

Fig. 22. The PixMix DR technique: on the left, the area to be removed is exhibited; on the center, the result from HERLING and BROLL [2010] approach; on the right the result from the PixMix technique from HERLING and BROLL [2012]

The Focus and Context technique (F+C) is another kind of visualization technique that is applied in AR/MR to deal with occlusion to define important areas that must be kept. It can be defined as a metaphor for visualizations to clearly differentiate very relevant information from the context [VIOLA et al., 2005]; examples of F+C techniques are: x-ray vision or ghosting techniques and cutaways [FEINER and SELIGMANN, 1992].

KALKOFEN et al. [2009] present a framework for interactive visualizations in AR. They exemplify in an efficient way the F+C technique. The content to be shown and removed is controlled through a technique called Magic Lens and different rendering styles can be used. This technique demands efforts in the preprocessing steps, as geometric modeling and classification of the objects.

MENDEZ and SCHMALSTIEG [2009] present importance masks to reveal information in x-ray; it is necessary to do preprocessing steps since the masks need the knowledge of the objects from the scene – an example of the use of their technique can be seen in Fig. 23. CHEN et al. [2009] do not need to do this kind of preprocessing procedures to get the geometry from the real scene in their F+C technique. The correct order and depth information between the real and virtual objects are computed with features, layers and mask obtained from video and virtual objects. But the lack of deep knowledge about the 3D scene is also a limitation to a better coherence.

Fig. 23. An example of the technique from MENDEZ and SCHMALSTIEG [2009] to manage occlusion

The technique from ZOLLMANN et al. [2010] builds a ghosting map to define the information on the scene to be preserved and removed (Fig. 24). A per-pixel- and a per-superpixel-based representation are used to analyze the image and this representation takes into consideration edges, saliency map, texture and synthetic details present in the scene; these data are combined to build the ghosting map. The user can also make adjustments in the transparency control. Despite using bottom-up factors, this is done as part of the identification of the important features, but not with the goal to direct the user attention. An improvement of this technique is presented in PADILHA et al. [2013]; the authors do an improvement in order to be able to work in indoor environments.

Fig. 24. Ghosting technique image based from ZOLLMANN et al. [2010]

A typical problem in x-ray vision is to define the kind of information to be presented. MR approaches deal with it in different ways. LI and NASHASHIBI [2011] present an x-ray vision to enable that the driver can see through the car in front of them; the technique utilizes GPS measurements and in order to get depth information, uses 3D perspective to transform based on 2D range perception. Further investigation on the use of stereo-vision is advised by the work. Besides selecting information, the filter algorithm from LIVINGSTON et al. [2011] manages occlusion and depth with the application of six metaphors: opacity, stipple, ground grid, edge map, virtual wall and virtual tunnel. Despite being used in the military mobile area, it can be used in different contexts.

An x-ray vision can be obtained too by the correct use and application of the concepts and theory related to transparency. Transparency plays an important role to realize the depth and occlusion in some cases. For a long time, artists have been using transparency to improve the perception of depth, occlusion and relations with others objects in scenes, as it can be seen in SAYIM and CAVANAGH [2011].

An interesting and important way to compute transparency is through X-junctions [SAYIM and CAVANAGH, 2011]. In the X-junctions, distinct objects overlap each other, producing an X-junction on the image [ADELSON and ANANDAN, 1990]. X-junctions' properties are used in the approach from FUKIAGE et al. [2012] to blend virtual objects in real scenes. The authors made psychophysical experiments to produce a model and consequently an algorithm to predict depth ordering in the blending process. Fig. 25 presents one result from their approach: on the left a simple insertion of the virtual object (house), and on the right the blending in agreement with their proposed model.

Fig. 25. Approach from FUKIAGE et al. [2012] to blend virtual object with real scene through the use of transparency theory concepts

TSUDA et al. [2005] define five visualization methods to show occluded objects in MR applications and to improve, in some methods, an x-ray visualization: grids, top-down view, and wire-frames are some of the resources used to improve the visualization of occluded objects; each method is an attempt to reduce ambiguity

KALKOFEN et al. [2013] present an adaptive ghosted views approach that improves the color contrast between occluding and occluded objects to improve the relationship between them.

3.1.4 DISCUSSION

In

Table 2, all the AR and MR approaches previously mentioned are summarized and organized by the visualization technique class that they take part in and the perceptual issue tackled. In agreement with this table, it is possible to define which kind of visualization can be applied to manage a specific perceptual issue.

Table 2. AR approaches organized by visualization technique and perceptual issue

Visualization Category	Perceptual Pipeline Problem	AR Approaches
Visual attention	Depth distortions and object ordering	SANDOR et al. [2010], AVERY et al. [2009], PADILHA and TEICHRIEB [2014], CHEN et al. [2011]
	Visibility	FISCHER et al. [2008], CHEN et al. [2011], WANG et al. [2010], HALLER and SPERL [2004], HALLER et al. [2005], FISCHER et al. [2005], MENDEZ et al. [2010], HANSEN et al. [2010]
Visual photometric consistency	Scene distortions and abstraction	NOH and SUNAR [2009], NOWROUZEZAHRAI et al. [2011], GIERLINGER et al. [2010], JACOBS and LOSCOS [2006], YEOH and ZHOU [2009], GRUBER et al. [2010], REINHARD et al. [2004], MENK and KOCH [2011], KNECHT et al. [2011], LIU et al. [2009], ZHU et al. [2010]
	Depth distortions and object ordering	MADSEN et al. [2003], SUGANO et al. [2003]
Visual geometric consistency	Visibility	ROLIM and TEICHRIEB [2012], LEAO et al. [2011], LIU et al. [2011], HERLING and BROLL [2010], HERLING and BROLL [2012]
	Depth distortions and object ordering	KALKOFEN et al. [2009], MENDEZ and SCHMALSTIEG [2009], CHEN et al. [2009], ZOLLMANN et al. [2010], PADILHA et al. [2013], LI and NASHASHIBI [2011], LIVINGSTON et al. [2011], FUKIAGE et al. [2012], TSUDA et al. [2005], KALKOFEN et al. [2013]

3.1.5 Lessons Learned

In agreement with the analysis of the visualization techniques to AR, we have realized that there is no pattern or rule to apply them. There is a tendency to apply x-ray or ghosting techniques to manage occlusion; some improvements on these techniques were done to manage other perceptions issues, as depth, as presented by FURMANSKI et al. [2002].

Emphasis on edges, color changes and animation are common techniques to get the user attention to important areas of the scene.

We have realized that Diminished Reality and Color Harmony techniques, in spite of being techniques already known, are not fully explored in the AR community as auxiliary techniques in the instruction process - we realized that they have great potential in this process. Section 4.1 presents our proposal in this sense.

3.2 Instructions in Augmented Reality Applications

Instructions are an important tool to help the usability of any kind of application; they are a way to give a feedback to the user and improve the learnability of the system [ROGERS et al. 2005]. Most works related with instructions have been concentrated in the assembly/manufacture area. We believe that these works are representative approaches to give an overview of instructions in AR and can be used in other application contexts too.

The effectiveness of AR in instruction-based systems has already been proven in some user studies, such as TANG et al. [2003], NILSSON and JOHANSSON [2008], NILSSON and JOHANSSON [2007], HENDERSON and FEINER [2009], ASAI et al. [2005] and WEBEL et al. [2013]. Here, we present an overview of how instructions are done in AR applications. It is easy to realize that there are two ways of giving instructions: the first one is the classical mode – the instruction is superimposed on the screen or image processing techniques are applied to emphasize parts of the scene to guide the user; on the second, the perceptual mode tries to manage perceptual problems before presenting the instruction.

3.2.1 Classical Instructions

The most basic and traditional way to present instructions is by text - it is simple and easy to understand. Other kinds of traditional instructions are: images, virtual objects and animations. The third person explanation is simple and often applied in AR applications; it can be seen as a union of virtual objects plus animation instruction. In this case, a virtual representation of a person tries to explain the steps to follow. For example, a virtual teacher explains/shows the instructions and the student/learner must perform the task at the same time; in cases where the instructions are related to body movements, the learner could do the same movements done by the virtual representation in a mirror-like interface [BLUM et al. 2012].

The work from BILLINGHURST et al. [2008] gives text and simple animation instructions on mobile phones. The application has two parts: the client side takes images from the scene and sends them to the server side to be processed to show the correct sequence and to create the animation; in their user studies, they concluded that AR plus animation was chosen by users as the best way to illustrate the sequence of steps in the assembly task. Their approach can be used in different contexts, because the main processing is done on the server, but it is necessary to put markers in the real scene. Instructions are superimposed onto the real scene, and occlusion between the real and virtual content is not dealt automatically. An example of their system can be seen in Fig. 26: the left image presents text and the start position of a virtual block, and the right image presents the final position of the virtual block.

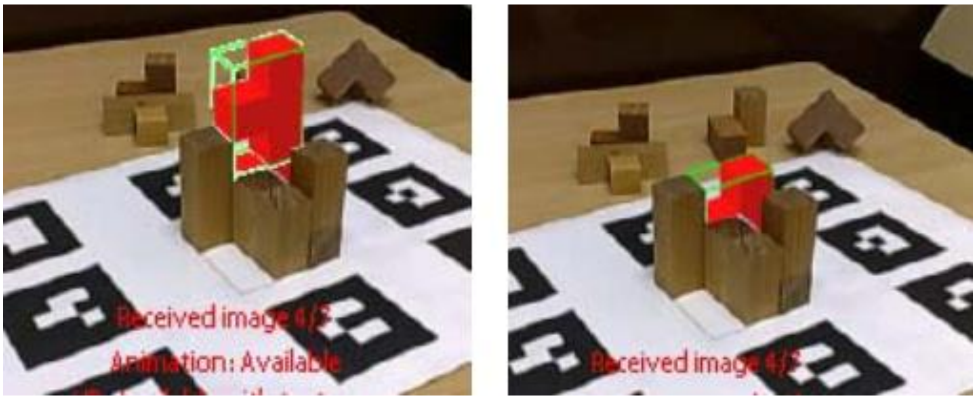

Fig. 26. An example of text and animation instructions from BILLINGHURST et al. [2008]

RAGHAVAN et al. [1999] give instructions by text, video, and 2D arrows on the screen in a marker-based system. Perception issues are not managed in their work, for example occlusion is one of perception issue that they intend to solve.

MOTOKAWA and SAITO [2007] present an approach that applies video instructions to help users play a guitar. Their system overlaps virtual 3D/2D objects to guide the user during the music.

In ANDERSEN et al. [2009] the guidance process of fixing a pump is done by text, video and images. Their approach have all the complete 3D CAD models of the real objects, they applied image features to estimate the pose in the real assembly sequence. Fig. 27 presents an example of their approach: the real object on the center must follow the steps indicated by images taken from the virtual representation of the object; these steps are presented on the top, and the correct position is presented on highlighted square.

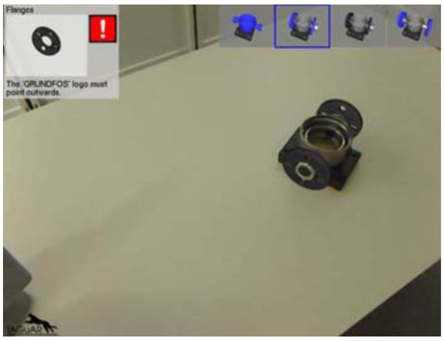

Fig. 27. An example of a session from ANDERSEN et al. [2009]; on the top, images indicate the next pose of the real object

HENDERSON and FEINER [2009] present a system for an armored personnel carrier turret; the instructions applied are texts, 3D and 2D arrows, and 3D virtual models of the tools (as presented in Fig. 28). ZAUNER et al. [2003] also apply text, images and 3D models to give instructions. An improved version of the approach from HENDERSON and FEINER [2009] is presented in HENDERSON and FEINER [2011] - they have dynamic 3D arrows and dynamic billboard labels. All instructions are superimposed on the scene. An example of dynamic arrows and labels is seen in Fig. 29: the labels (letter J in blue square) change position in agreement with the user view, furthermore the arrow changes color and direction if the task is performed correctly or not.

Fig. 28. Example of use of the system proposed by HENDERSON and FEINER [2009]: instructions are done with texts and virtual 3D models

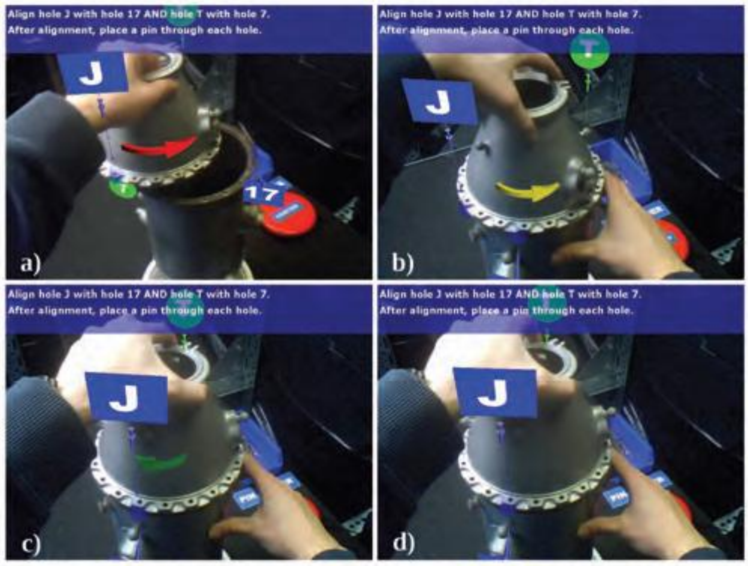

Fig. 29. Example of dynamic arrow from HENDERSON and FEINER [2011]: (a) large, red, dynamic arrow indicates direction and magnitude of motion needed to align can and cone; (b) as can and cone approach alignment, arrow reduces size and changes color from yellow to green; (c) it also alters direction to specify shortest rotational direction to alignment; (d) when alignment is achieved, arrow fades away

Some works superimpose a virtual agent on the screen to assist the user, as we have mentioned before. Major examples are the approaches from MIYAWAKI and SANO [2008], NAWAHDAH and INOUE [2011] and WANG et al. [2013]. Their approaches are marker based and don't manage perceptual issues, as occlusion or depth perception. An example of virtual agent is presented in Fig. 30.

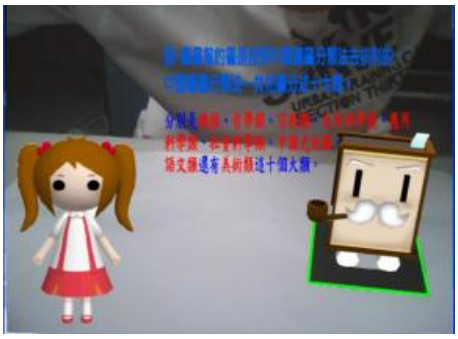

Fig. 30. A virtual agent from WANG et al. [2013]

MARCINČIN et al. [2011] apply 3D models of tools in different colors, created previously, to aid the user. A use of their system can be seen in Fig. 31. The connection between the real and virtual content is done by sensor and image processing techniques. We believe that the possibility to use their system in different scenarios would be a challenge because of the creation of 3D models and the use of sensors.

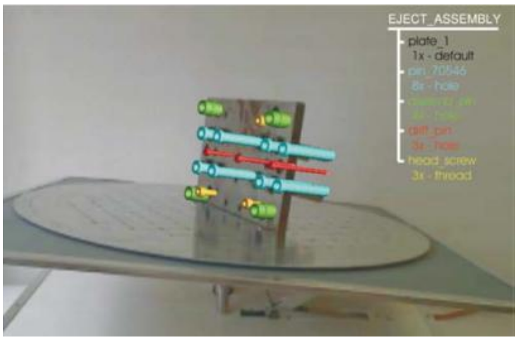

Fig. 31. Instructions based on 3D models of tools in different colors [MARCINČIN et al. 2011]

PATHOMAREE and CHAROENSEANG [2005] apply text and image processing techniques (edges identification to emphasize parts to be moved or the correct place to move) to guide and help the users. The tracking method is done by a CAMSHIFT algorithm [BRADSKI, 1998].

GOTO et al. [2010] apply preexisting 2D videos as instructional videos, in agreement with the user view. The video information is superimposed onto the user view, after the application of image processing techniques, such as edges extraction and x-ray vision – this is important to emphasize parts of the image and/or define the level of transparency in x-ray vision. Fig. 32 presents an example a session of their approach. The instructor's hand is an x-ray visualization created after processing the frames got from the user view and the video recorded previously about the task. They utilize markers to identify the user actions. PETERSEN et al. [2013] use a similar approach in the guiding process, with half-transparent overlays and colors, but do not apply markers to identify user actions - they use image processing techniques and 2D probability maps. Whilst the task segmentation in GOTO et al. [2010] is made mostly by hand, the approach from PETERSEN and STRICKER [2012] presents a method to automatically segment the task in the instructional video, by creating instructional overlays during task performance.

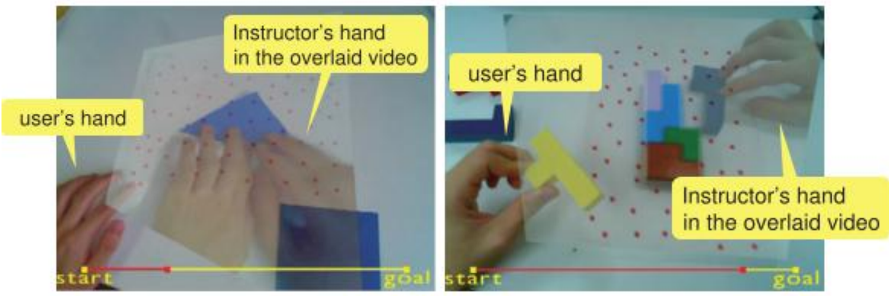

Fig. 32. Instruction process done with x-ray vision from the approach by GOTO et al. [2010]:

the instructor's hand, from an instructional video, is superimposed in real-time on the user's hand with a defined level of transparency

WEBEL et al. [2013] classify the rendered instruction information in two kinds: *Direct Visual Aid*, where the 3D information is directly superimposed onto the captured video image; and *Indirect Visual Aid,* that indicates what contextual information is available displaying the machine part. To validate their proposal, they present a system whose instructions are done with text, videos and 3D models (some animated others not). The authors don't explain how the information is created, saved and edited to be used in other contexts. The tracking is done with vibrotactile bracelets; they allow recognizing specific movements, as rotation.

HORIE et al. [2006] propose an interactive cooking learning application based on a MR environment. The instruction tips used are 2D arrows to indicate the destination of an ingredient (Fig. 33, left image), a spiral indicating how to (the direction of mixing) mix ingredients (Fig. 33, right image), and text messages to indicate if the food is hot enough (they apply an infrared camera to identify the food temperature), and videos about cooking. The tracking and identification of action are done by infrared cameras and markers.

Fig. 33. Instructions in a cooking AR application from HORIE et al. [2006]

3.2.2 PERCEPTUAL MODE

This category presents the approaches that deal with perception issues to present instructions to users. Instructions are basically the same as the classical ones, but some processing is done to deal with depth and occlusion, for example.

GIMENO et al. [2013] present SUGAR, a framework to enable rapid prototyping of low-cost AR systems. SUGAR has two parts, one to get the real-world description and another to edit the virtual information. The real-world description is done through real images from the scenes as well as 3D scenario created by a RGB-D sensor. Through this 3D scenario it is possible to manage depth issues to place correctly the instruction. The virtual content is located in a repository, but there are basic 3D objects, as arrows to use to create instructions. The user must have some knowledge about 3D edition to edit/use/create these models.

REINERS et al. [1998] present a complete system to fix a door lock - all the system, instruction and model was specific to this task. The 3D virtual models used were created previously. To identify the actions and real objects movements, optical markers were applied. The occlusion was managed by a Z-buffer approach [HUGHES et al. 2013]. Instructions were done with animations from videos created previously too and 3D virtual models. Fig. 34 presents an example from their approach: virtual hand and object (pink color) are presented in agreement with depth of the scene and occlusion situations.

Fig. 34. An example from the approach from REINERS et al. [1998]: occlusion and depth perception issues are manage to improved the task to fixing a door lock

The approaches from KHUONG et al. [2014] and GUPTA et al. [2012] are developed to perform LEGO® based tasks. The first work proposes two illustration ways to give the instructions. In the first, the assembly instructions show the next blocks to add as 3D animated wireframe, directly overlaying the physical model. In the second mode, they called side-by-side, the assembly instructions are displayed on top of a separate virtual model that is rendered beside the actual model (Fig. 35). Previous reconstruction and tracking are done by an algorithm called Lattice-First [MILLER et al. 2012] with the Kinect RGB-D sensor [KINECT, 2012] (the depth information is very important to the success of the algorithm). Their reconstruction and tracking are specific to building blocks, so the use of it in different contexts is not possible.

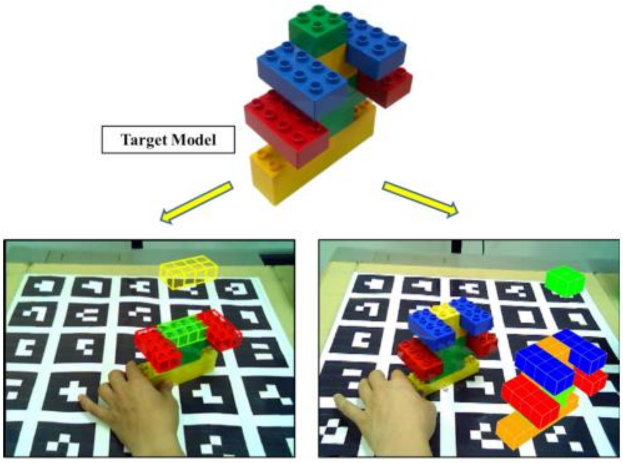

Fig. 35. Two proposed instructions modes from KHUONG et al. [2014]: on the left, wireframe, on the right, side-by-side model

GUPTA et al. [2012] guide the user through the application displaying 3D virtual content on the screen. Tracking and guiding are done by a combination of color-based and depth-based approaches. There are two kinds of mode: authoring and guidance. In the authoring mode the user can reconstruct a virtual model from a block, add or remove parts. In the guidance mode the models created in the authoring mode are used to create the sequence. Their approach works only with building blocks.

IWAI et al. [2013] propose a projection-based approach that applies an energy function to the correct placement and depth perception of the instructions presented. Fig. 36 illustrates their approach where in agreement with the user's viewpoint, the energy function defines the correct shape and placement of the instructions (ABC letters on the face of the dummy).

Fig. 36. Projection-based AR instruction from IWAI et al. [2013]

Another way to give contextual instructions in AR is through the use of annotations. WITHER et al. [2009] define annotations in the AR context as virtual information that describes in some way, and is registered to, an existing object.

This definition is wide, and allows its use in different kinds of AR scenarios. The placement of annotations must deal with the perception of depth, overlapping and to be sure that important objects are not obscured. The authors define six orthogonal dimensions to describe annotations: location complexity, location movement, semantic relevance, content complexity, interactivity, and annotation permanence. MAKITA et al. [2009] present an approach that creates a probability map and penalty scores to put the annotation in the correct position (avoiding occlusion between the elements). Their goal is to apply the approach to targets/objects that are moving or are non-rigid. An example of their approach can be seen in Fig. 37.

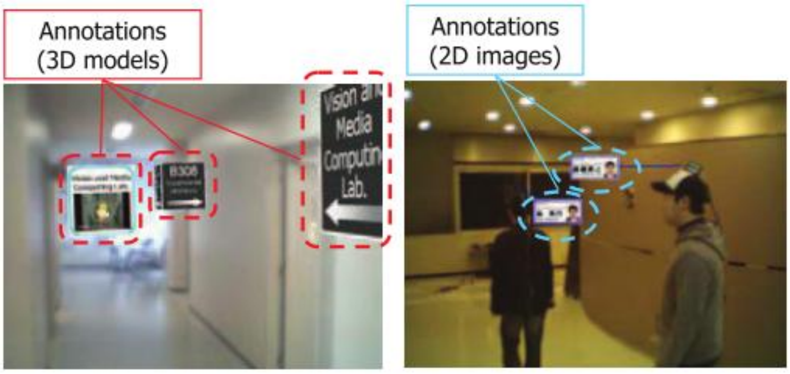

Fig. 37. Instructions given by 3D objects (left) and images (right) located correctly in agreement with the approach from MAKITA et al. [2009]

URATANI et al. [2005] utilize monocular depth cues (for example, texture gradients and relative size) to change the position and appearance of the annotation. They applied images and texts in their approach.

TATZGERN et al. [2013] apply filters and a layout algorithm to organize the annotation when there is a lot of information to visualize, mainly in small screens, as mobile devices. Their kinds of annotations are textual and 2D/3D objects.

3.2.3 Body Instructions

All of the previous works are related to object instructions. However, few works are related to body movement instructions in AR. An example is the work from GAMA et al. [2012a]. Their approach guides and corrects user movements with text instructions. They utilize a RGB-D sensor to capture user actions and a planes and checkpoints technique to recognize the movement and verify if it is correct according to the movements preregistered to the system. They applied texts (Fig. 38, left image) and objects (Fig. 38, right image – top and down) to guide and give feedback to the users.

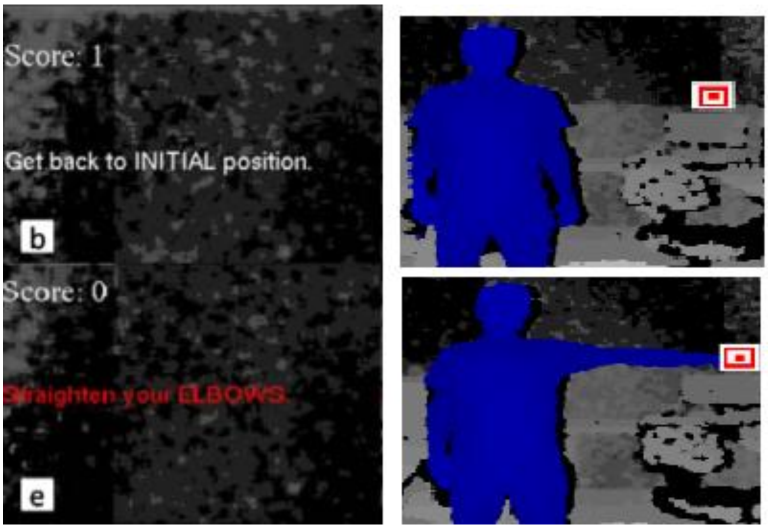

Fig. 38. Body instructions from the approach by GAMA et al. [2012]: on the left text instructions; on the right (top and down) an object guides the movement

YouMove is an approach proposed by ANDERSON et al. [2013] to teach previously recorded physical movements. It applies an RGB-D sensor to record and capture the user movements. The instructions are presented by the exhibition of texts, videos, images and teacher's skeleton saved previously – the last one is the only way that the users have the correction and guidance to their actions, as it can be seen in Fig. 39.

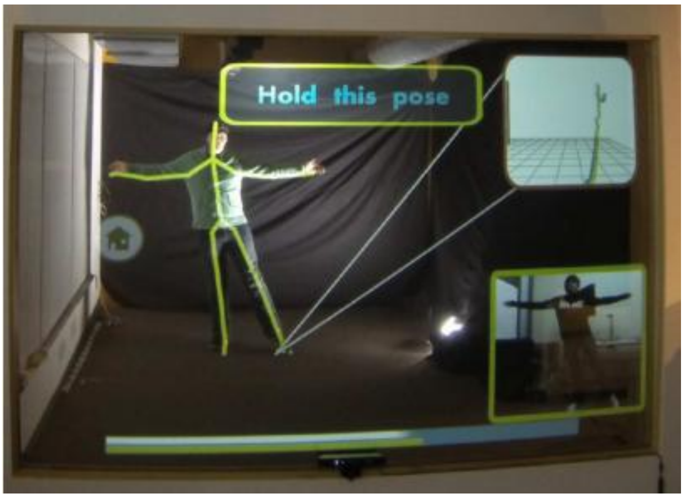

Fig. 39. Example of instruction from ANDERSON et al. [2013]

Another approach related is the Physio@Home from TANG et al. [2015]. This approach gives body instructions based on the information received from sensors and markers, as shown in Fig. 40. Despite considering the sensor data to present the instructions, users have difficulty to realize and understand the depth, as informed by the authors.

Fig. 40: Example of instruction from TANG et al. [2015]

3.2.4 Discussion

Table 3 presents a summary of the AR/MR approaches discussed in this Chapter. Its main goal is to present the main types of instructions applied in AR, as well as the way in which the instruction is presented: a simple superimposed content on the real environment or using image processing techniques to blend the augmented content with the real one in order to enhance user perception.

Table 3. A summary of AR classical instructions approaches and how they are presented to the user

Instructions Presentation	Types of Instructions	AR Approaches
Superimposed	Text/Animation	BILLINGHURST et al. [2008]
Superimposed	Text/Animation/2D arrows	RAGHAVAN et al. [1999]
Superimposed	Text/3D arrows	HENDERSON and FEINER [2011]
Superimposed	Text/2D and 3D arrows/3D models	HENDERSON and FEINER [2009], WEBEL et al. [2013]
Superimposed	Virtual agent	MIYAWAKI and SANO [2008], NAWAHDAH and INOUE [2011], WANG et al. [2013]
Superimposed and procedure-ray vision	Text/3D models/Images/Video	ANDERSEN et al. [2009], GOTO et al. [2010], PETERSEN and STRICKER [2012]
Superimposed	Text/Images	PATHOMAREE and CHAROENSEANG [2005]
Superimposed	Text/3D models	MARCINČIN et al. [2011]
Superimposed	Text/2D models	GAMA et al. [2012]
Half-transparent overlays and colors	Video	PETERSEN et al. [2013]
Superimposed	Text/Animation/3D objects	ZAUNER et al. [2003]
Superimposed	Text/2D objects/Video	HORIE et al. [2006]
Superimposed	2D/3D objects	MOTOKAWA and SAITO [2007]

All the approaches that we found related to the perceptual mode manage depth distortion and object ordering. They are presented in Table 4. We also specify type of instructions applied and the approach applied to manage the perceptual issue.

Table 4. AR approaches that manage some perceptual problem to create the AR scene

Types of Instructions	Way to Manage the Perception Problem	AR Approaches
Text/Images/3D models/3D arrows	Depth map of the scene	GIMENO et al. [2013]
Video/3D objects	Z-buffer approach	REINERS et al. [1998]
Animation/3D objects	Depth map of the scene	KHUONG et al. [2014], GUPTA et al. [2012]
Text/Images	Energy function	IWAI et al. [2013]
Images/3D objects	Probability map and penalties scores	MAKITA et al. [2009]
Text/Images	Monocular depth cues	URATANI et al. [2005]
Text/2D objects/3D objects	Filters and a layout algorithm	TATZGERN et al. [2013]
Texts/2D objects/3D objects	Depth map	ANDERSON et. al [2013]
Texts/2D objects/3D objects	Sensor and markers	TANG et al. [2015]

The use of text for instructing is common in AR applications, but with tasks that deal with movements, it is usually not enough. Few approaches manage perception problems before presenting the information, besides most of them are developed to a specific environment/task (for example to fix a door, to fix a car engine), so the possibility to apply the same AR application in different environments/tasks is limited. A solution is the use of an authoring framework, e.g. the framework SUGAR developed by GIMENO et al. [2013], but non-expert users can have difficulty to use it.

It can be observed that most of instructions approaches used in the literature are basically the same: information presented using text, images, videos, 2D and 3D objects; each work defines its own set of instructions according to the task focused in the application. Another observation is that the use of instructions related to body movements is still to be explored – we found only one work in the literature.

3.2.5 Lessons Learned

The typical kinds of instructions applied in AR applications are: text, animation, 2D/3D objects, color changes, edges emphasize, images, videos. We can realize from Table 3 that, usually, most of the AR approaches applies at maximum three of them. This happens maybe because of the difficulty to implement them or maybe because the developers hope that the instructions implemented were enough; but it would be interesting for them to count with guidelines that they can follow when there is the necessity to give instructions to users.

We realized that most of the AR approaches focus on object instructions and few works in body instructions. However most of kinds of instructions can be used in both cases; so, our guidelines have the goal to be general to be applied to object and body instructions.

3.3 Design Guidelines for Generating Augmented Reality Instructions

Whilst AGRAWALA et al. [2011] defined their guidelines based on Cognitive Psychology research, NIENHAUS and DÖLLNER [2005] extract principles for illustrating motion from concepts found in comic books and storybooks. Both sources only consider purely virtual 2D and 3D illustrations, and must be applied to AR with caution.

Some principles from AGRAWALA et al. [2011] and NIENHAUS and DOLLNER [2005] are common to both approaches, and some are concerning AR only. So, this thesis we have surveyed the most representative approaches related to visualization and instructions in AR (sections 3.1 and 3.2); we analyzed what kind of instructions were used and how they manage the perception problem.

The most popular means to visually guide the user in the task are text, images and animated 3D models. These very conservative choices indicate that the potential of AR visualization techniques is often not fully realized.

Therefore, in order to give instructions in AR based systems, this thesis proposes the following guidelines [ROLIM et al., 2015]:

(1) *Indicate movement:*
An instruction must indicate the correct path, the correctness of the movement and, in some cases, the velocity or acceleration. The path sets the trajectory of the movement whereas the correctness indicates the right way to achieve the goal ('is it better to go left or right?', 'Which previous movements do I need to do?'). With some kinds of instructions, the user must follow some velocity to reach the goal. The approaches related in the classical instructions can be used here.

(2) *Emphasize parts of an object or a body to be moved or changed or get the user attention:*

It is important for body instructions to guide the user about which part of the body must be moved. In object instructions, when there is a lot of small objects that composite a big one, it is important to identify which one must be moved or to have the direction changed. In this guideline, the AR/MR visualization techniques from the visual attention category, which are related to the visibility perceptual problem, and visual geometric consistency category, which are related to the perceptual problem *scene distortions and abstraction*, can be used to focus the user attention and to identify objects in scene.

(3) Allow different kinds of visual appearance attributes:

The visual appearance of AR instructions should agree with the environment conditions (light, pattern of colors and so on). The user should remain in control over the appearance of the instruction?, because some tasks may have specific patterns to follow. The approaches related in the classical instructions can be used here.

(4) *Feedback*:

The process of giving an instruction must deal with feedback for users to convey if the user is proceeding correctly, what must be changed, or if there are alternatives. It is essential that such suggestions are presented in real-time.

(5) *Management of occlusion and depth:*

Instructions conveying three-dimensional information, such as a disassembly task, must incorporate visual cues to let the user understand occlusion, depth relationships and distances in general. In the application of these guidelines, we must take care that perceptual problems are properly managed. The AR/MR approaches, from our classification in section 3.1, which deals the perceptual problem *depth distortions and object ordering*, can be used here to try to manage the occlusion and depth issues.

3.4 CONCLUSION

We did a survey about the most representative visualization and instructions techniques with the aim to verify when and how they are applied in AR applications. From this analysis, lessons learned and related works, we could identify patterns and procedures to propose our guidelines. We had an initial set of guidelines to be implemented and validated; besides, we had a proposal to be criticized in the sense of being improved by the AR community.

4 INSTANTIATION STAGE

In this stage, we implemented the design guidelines. We analyzed how to apply AR visualization in an instruction process; we presented our system to create automatic instructions in agreement with our guidelines and some results produced by the usage of the system.

4.1 VISUALIZATION TECHNIQUES APPLIED TO AR

For each visualization category presented in Chapter 3, we found a representative approach to indicate how to apply it in instructions in AR:

- In visual attention: the ghosting technique was chosen in this category because it can be used to keep the context available to the user during the execution of a task. To illustrate this application, Fig. 41 presents an approach from KALKOFEN et al. [2011] that gives information to user (the content of inside the car), so he/she can realize the context (car) during a task – for example, when changing some pieces in the inside of a car, which parts can/will be affected.

Fig. 41. Ghosting example in AR

- In visual photometric consistency: the technique Color Harmony is applied to remove the user attention from virtual objects. So, when the task demands that the user changes their focus of attention from a virtual object, this technique could be applied. Fig. 42 is adapted from WONG et al. [2012]; the left image illustrates a virtual car added to the real scene, its color is different from the scene context, so it is highlighted from the other parts, and this means that the user has the tendency to look first to it; when the color harmony approach is applied, on the center image, the user can realize other parts of the scene, and other virtual objects could be added and direct the user attention – on the right a virtual arrow was added to direct the user attention to the window.

Fig. 42. Example of color harmony approach applied to AR instruction

- In visual geometric consistency: Diminished Reality could be applied to remove undesirable areas and keep the user attention focused to important ones, as it can be visualized in Fig. 43: on the right, a panel from an airplane with a lot of areas to learn. A user in training find it difficult to understand the panel at first, so the idea is to remove parts of the panel (center and right images) for the trainer to focus on specific areas at a time and perform the activity.

Fig. 43. Example of Diminished Reality applied to AR instruction

We worked in DR approach based Jan Herling and Wolfgang Broll [HERLING and BROLL, W. 2010]. This approach does not need any preprocessing steps like background information. But it is necessary to identify/select the object on the video image. It is possible to use the approach with objects and people. Two results are presented in Fig. 44 – the top image illustrates on the left some objects on a chair that are removed on the right image; on bottom image illustrates the removal of an undesirable person.

Fig. 44. Results of our developed DR approach

4.2 Guidelines to AR

From the analysis of the application of visualization techniques and instructions in AR applications, for each guideline that was proposed in Chapter 3, we proposed and implemented prototypical tools for generating instructions and validate this proposal. Table 5 presents the kind of instructions that can be applied in each proposed guideline. We implemented some of the instructions illustrated in this table, and they are explained in details in next section.

Table 5. Guidelines and instructions proposed

Guideline	Instruction proposed
Indicate movement	Arrows, stroboscopic motion, key poses, images, animation
Emphasize parts of an object or a body to be moved or changed or get the user attention	Edge outlines, diminished reality, animation, color harmony
Allow different kinds of visual appearance attributes	Color changes, 2D/3D objects, color harmony, animation
Feedback	Text, color changes, arrows
Management of occlusion and depth	Avatar, ghosting techniques/x-ray vision, key poses, arrows

4.3 Augmented Reality Instruction System

As it was discussed in the previous chapter, this chapter explains the AR instructions system developed in the instantiation stage to validate this proposal. Different instructions were implemented to be tested and validated in the test scenarios.

4.3.1 INFORMATION FLOW

Fig 45 illustrates the basic information flow to use the system for presenting body instructions. At first, an expert does the sequence of movements or a pose that could be done by others. The system records all the available information, as RGB data, depth data and skeleton data – it is like having a database of poses (it is illustrated by two red skeletons inside of a black square). The normal time to reach the pose is recorded by the expert.

A set of instructions is implemented to be used and chosen by the users (represented in Fig. 45 by the straight arrow, curved arrow and the word text inside of a black square). The user, before doing a task, chooses the kind and the level of instruction (high or low). After that, the user does the movement in accordance with the instructions presented. The system checks the pose tracked (blue skeleton) against the skeleton saved (red skeleton); if they match the task is considered complete, and if they do not match an instruction is presented correcting the movement or indicating the correct movement. The task is complete when the skeletons match or the time is reached.

It is possible to realize that the system allows the use in different kinds of situations (body workouts, yoga classes, dance and so on), because it is not related to a specific set of poses, movements or tasks.

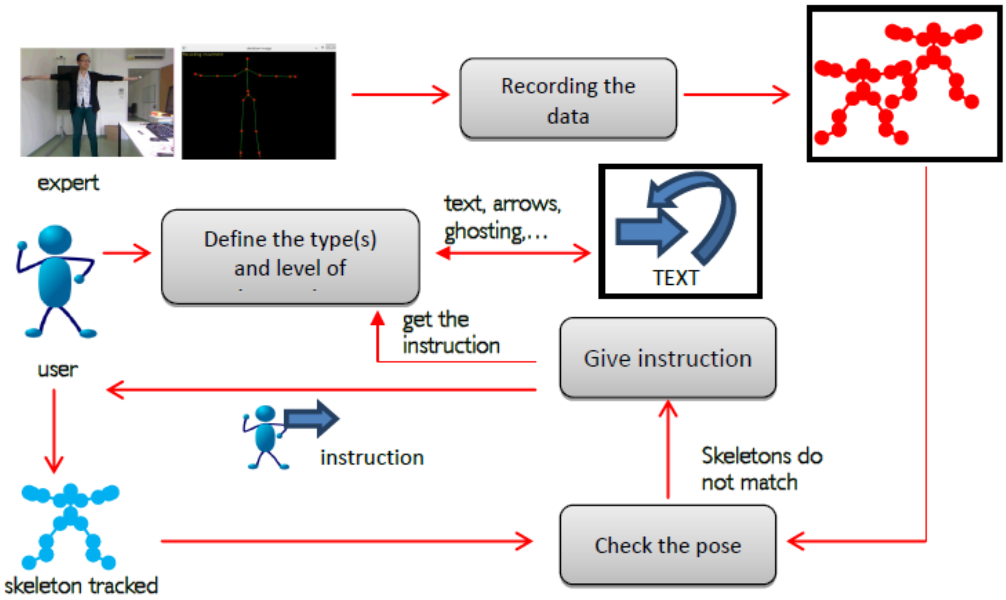

Fig. 45. Illustration of the basic flow of information in the system

4.3.2 *SYSTEM DESCRIPTION*

An overview of this system is presented in Fig. 46. A front-end module called façade module is responsible for keeping an interface between the other modules and various AR applications. To save a sequence of poses demonstrated by an expert, we use the Kinect SDK [KINECT, 2012] in a recording module. We record depth, skeleton data, images and video. The illustration module is responsible to generate instructions, including arrows, edges, avatar, texts, key poses, and ghosting views. This module interacts with the recognition module to check if the user's movements are correct with respect to the reference movement performed by an expert. The reporting module is responsible to give an overview of the user's actions. The arrows indicate the flow of information.

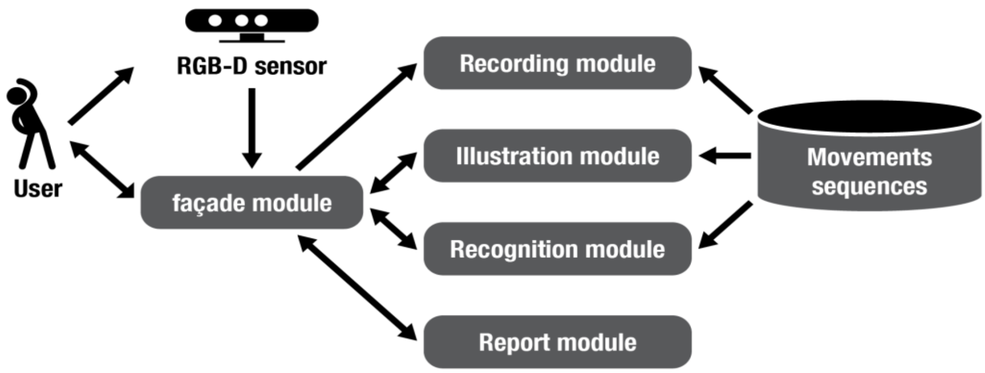

Fig. 46. The architecture of the AR instructions system

The 'movement sequences' module, presented in Fig. 46, is a way to illustrate that the poses and movements sequence are saved to compare with the actual user pose to verify if the user is doing right or wrong.

The system is implemented in C++ [DEITEL and DEITEL, 2008], and the illustration module uses the libraries OpenCV [BRADSKI and KAEHLER, 2008] and OpenGL [SHREINER et al. 2013] to generate instructions. The avatar instruction, which is part of the illustration module, is implemented with C# [DEITEL and DEITEL, 2003] and the XNA framework [XNA GAME STUDIO, 2015]. There are 3D arrows that were created in 3D Studio Max software [AUTODESK, 2015].

We have decided to implement the system in modules because it is easier to change or improve specific modules for better solutions. For example, there are approaches in the literature which work only to recognize and record the user movements, as presented in CHAVES et al. [2012] – actually, our recognition module is simple version of their work, to more details check their approach. To check if a pose was reached the system defines a set of vectors from the tracked skeleton and from the saved pose – each arrow in Fig. 47 indicates a generated vector.

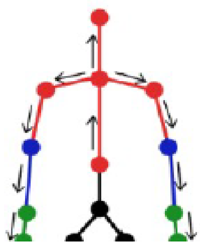

Fig. 47. The arrows indicate the vectors defined by the AR instructions system

The system calculates the angle that exists between each correspondent vector from the saved pose and from the tracked skeleton: if the angle is equal or less to a pre-defined angle, the pose is reached. This process is illustrated in Fig. 48: the left image shows the tracked skeleton of a user; the vector defined by the forearm and the left hand is compared to the vector from the same body parts of the skeleton saved previously when an expert user performed the specific movement correctly. This correct movement can then be compared during the use of the system by any user in order to verify correctness.

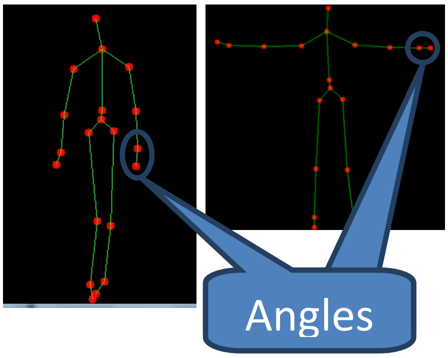

Fig. 48. Example of angle pose checking: left image presents an illustration of a skeleton tracked in real-time, right image presents a skeleton saved in our system. The angle defined by the vector between the hand and forearm from the tracked skeleton is compared with the angle of the same vector from the saved pose

4.3.3 RESULTS

This section presents results related to body and object instructions.

4.3.3.1 BODY INSTRUCTIONS

Fig. 49 presents a session recording of a sequence of body movements: on the left image, the RGB view of the user performing the movement is shown whereas on the right image the skeleton data of the pose captured by the Kinect sensor at each frame of the motion sequence is presented.

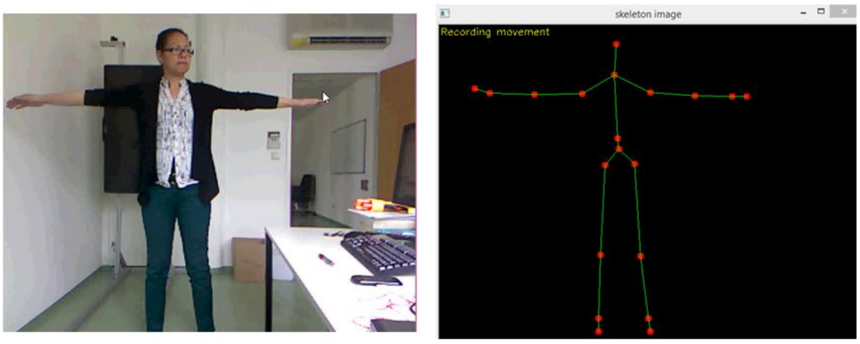

Fig. 49. Example of a body movement recording with the AR instructions system

Next, we will present some results, in agreement with the guidelines defined.

4.3.3.1.1 EMPHASIZE PARTS

The users can identify the parts of the body to be moved by the use of edges (see an example in Fig. 50). Edge extraction is a common technique applied in image processing techniques to highlight areas from images. In this thesis, we also applied it to highlight areas of images but in the sense of giving an instruction.

Fig. 50. Example of edges and arrows to indicate the part of the body to be moved and the direction of the movement

The system combines the information from depth, skeleton and RGB to create the edges instructions. To speed up the processing time, the system defines first a region of interest (ROI) on the images – a rectangle area containing the joints or body parts that the user wants to track. The ROI limits are defined by the skeleton data. After defining the ROI, a smooth filter is applied on the depth image to remove small and unimportant details that can be present on the user environment. The Canny algorithm [CANNY 1986] is used to identify the body contours, followed by a dilation technique, and this information is compared with the skeleton data, to highlight the body contours in the RGB image. An example of this process is illustrated in Fig. 51. The parameters from the smooth filter and Canny technique are easily configurable. Improvements must be done in edges based instructions when the user is in a noisy background, because the system cannot differentiate them – a Diminished Reality technique can be applied to remove undesired areas.

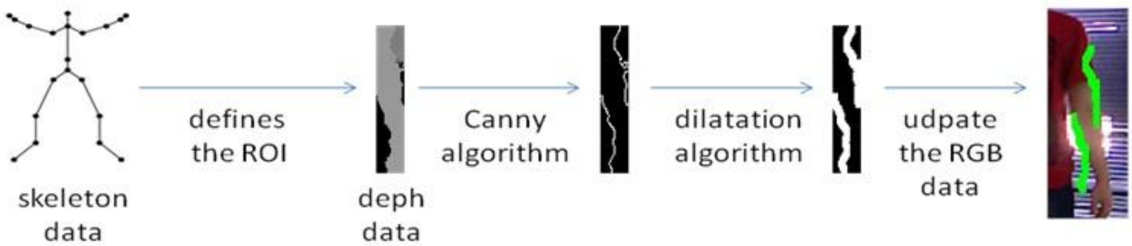

Fig. 51. Steps of the creation of the edges instruction

4.3.3.1.2 Indication of the Direction of the Movement

The arrows are the simplest visual way to indicate the direction of any kind of movement and frequently applied in AR applications (see chapter 3). The system allows the use of two kinds of arrows: straight and curved. Both indicate direction, but with different goals; since the movement can be composed of a set of poses, the straight arrow receives the actual user pose and indicates the next pose until the user completes the full movement.

The curved arrow receives all the poses that comprehend the recorded movement and tries to illustrate the path of the movement that the user needs to follow. The user has a full view of the movement before trying to do the movement. The points that define the curved arrow belong to the body part that the user wants to illustrate in the instruction.

Fig. 52 is an adaptation of GAMA et al. [2012a] and it illustrates, on the left, an example of using of straight arrow; the direction is done for two poses that are part of the full movement; the image on the center illustrates the green points that were got from all the poses to define the curved arrow presented on the right.

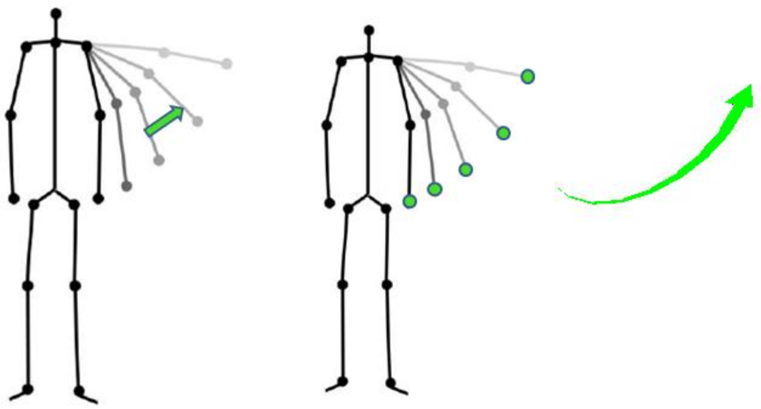

Fig. 52. Illustration of the process to generate the straight and curved arrows

4.3.3.1.3 MANAGEMENT OF OCCLUSION AND DEPTH

The user can see the movement or a saved pose with an avatar, as illustrated in Fig. 53 - on the left, real images that represent saved poses, and the illustration with an avatar on the right. To create the ghosting instruction illustration, a blending of the real images from the user and images from the avatar is done (top left image in Fig. 53). The avatar model applied in our tests is a model available in the C# Samples from Kinect SDK [KINECT, 2012].

Fig. 53. Example of illustration of the movement with an avatar

Besides visualizing the poses that comprehend the movement, the user can visualize only the next pose to execute: the mirror view is divided into two: one presents the avatar in the right position, and the other the view of herself/himself, as it is illustrated in Fig. 54.

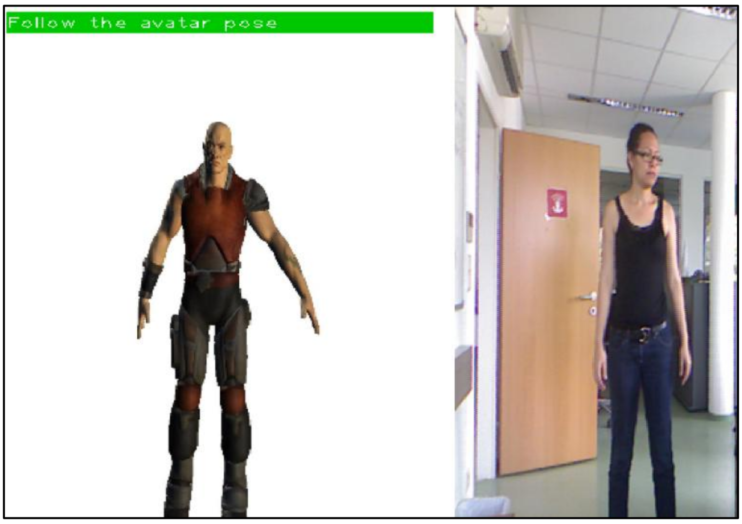

Fig. 54. The avatar on the left illustrates the correct pose to the user on the right

4.3.3.1.4 Giving Feedback

Messages and 2D drawings are used as a way to help the users and guide them during the movement. This is illustrated in Fig. 55: the text on the top of the screen informs general messages, the red line informs the wrong position, the green line informs the right position. The red line has the goal to illustrate the actual pose of the user skeleton (part of the skeleton, the figure is related to the left arm). The green line tries to inform the user the exact position to put the skeleton. At the bottom of the image is illustrated the percentage of the movement reached. Some tips are presented to help the user to do the movement right.

Fig. 55. Example of feedback to users

The system allows the user to choose the level of instruction related to body movements as high or low. Low level means the correct direction/movement is shown for each bone, while high level shows instructions for a set of bones in an aggregated way. The set of bones for each level is defined in agreement with the skeleton data from the Kinect API [KINECT, 2012] (see Fig. 56). Fig. 57 illustrates on the left the instructions by arrows in low level, and on the right, in high level.

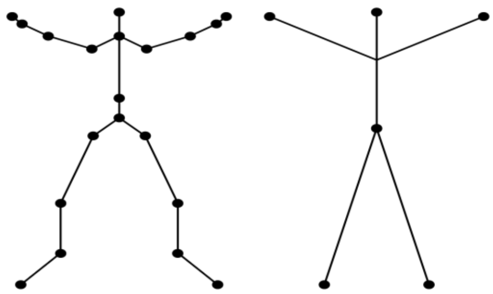

Fig. 56. The levels of bones used in the instruction process: on the left, low level; on the right, high level

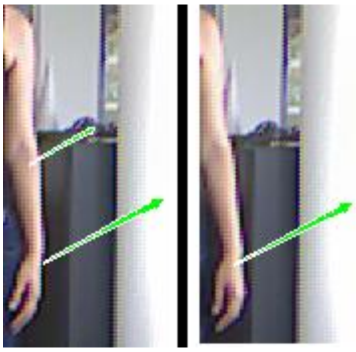

Fig. 57. Levels of instructions: low level (left), high level (right). The instructions are represented by straight arrows

4.3.3.2 Object Instructions

Despite the fact that the previous results are illustrated with body movements, some of them can be applied with object instructions. This section presents results and specific issues related to object related instructions. These results are based on information from a RGB-D sensor, to create the instructions.

4.3.3.2.1 Indication of the Direction of the Movement

IZADI et al. [2011] developed the KinectFusion, that allows the 3D reconstruction of a scene through data collected with a RGB-D sensor, such as the Kinect. We applied their API (Application Program Interface) to create a 3D model of the scene, and identify changes in this model. Through the identification of changes, it is possible to identify the path of objects.

Fig. 58 illustrates the main steps of object instruction: first the scene is reconstructed to create a 3D model; if any change happens to the model, the system detects it and saves the 3D position of the areas changed. So, it is possible to follow the path of these changes to create instructions to objects (center and right images).

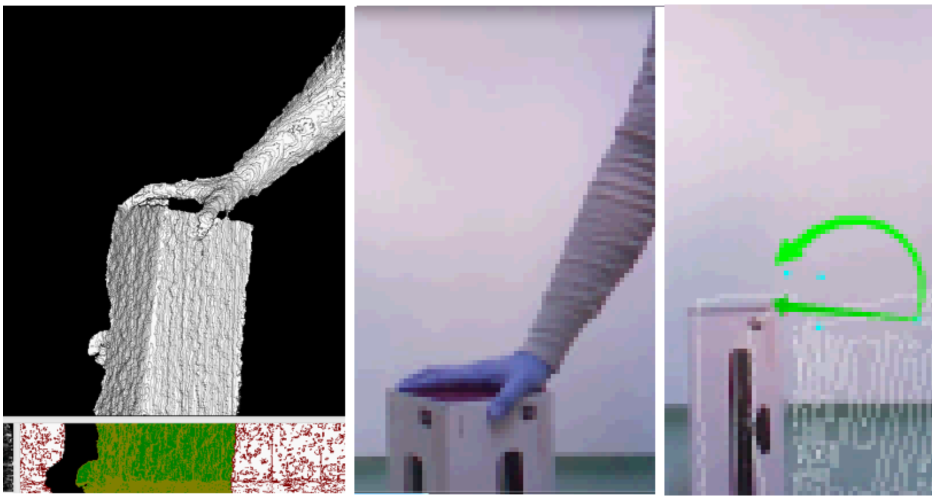

Fig. 58. Steps to create object instructions based on a 3D model of the scene

4.3.3.2.2 Management of Occlusion and Depth

An instruction is a virtual information that must be able to deal with real objects to correct placement of it. Fig. 59 presents a virtual rectangle that goes from the left side to the right side of the scene, but the real red wallet must be in front of it when both are at the same position. The priority of the wallet is defined by a mask defined by the user, in agreement with the depth data obtained from the scene. The algorithm verifies for each pixel of the virtual box if they match with the area from the mask.

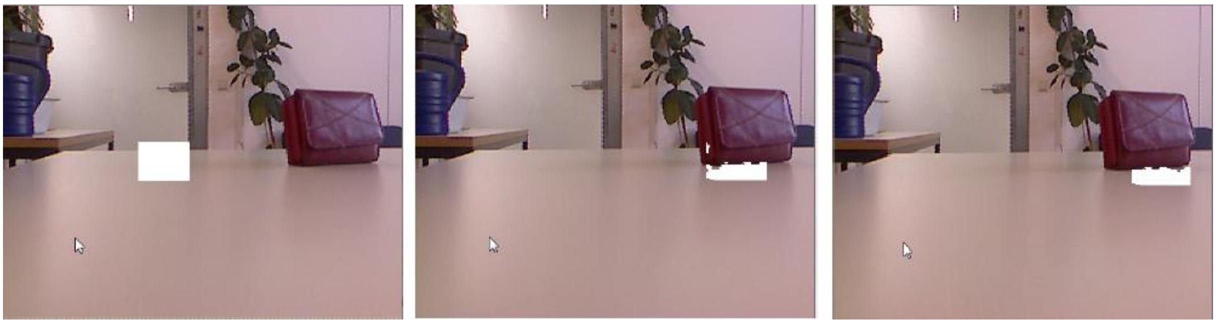

Fig. 59. Three different positions of the virtual object to present the occlusion effect

The ghosting effect is interesting procedure applied to visualize information or objects behind real objects. So, this effect could be used by an instruction, to help the user understand scene context and to do correct the task. The ghosting effect is illustrated in Fig. 60. The frames from the videos (with the real object and without the object) and the 3D model of the wallet were used to compose the final image based on the alpha blending process. In this process, the color information is combined between two images to generate the transparency effect.

The frame from the 3D wallet was used to define the parts of the scene that can be seen as transparent. The alpha value was set from 1 (the real object is totally opaque) to 0 (totally hidden). This algorithm has a limitation related to the correct registration of the 3D object on the scene, and because of this, the real object has problems on the borders.

Fig. 60. Example of ghosting on real objects

Another approach developed creates an importance map from the real scene in order to blend in the virtual objects in a more adequate way. A careless augmentation can prejudice the user AR based experience. The importance map is built from a saliency map, edge detection, superpixel representation and texture analysis. Fig. 61 illustrates two results that were obtained from our approach: on the left the virtual green pipes are presented in agreement with the importance map created, so that the context is preserved and the user has better information about the relation between the virtual and real objects. On the right, a virtual chair is placed on a real scene. This kind of approach could be applied to the automatic identification of important elements that are used by create an instruction.

Fig. 61. Results from ghosting AR approach developed

4.4 Conclusion

We presented an AR system with a set of instructions to be applied to body movements. This system was developed in agreement with the guidelines proposed and it will be used to validate the instructions and visualization techniques by end users. Besides, we illustrate how common visualization techniques can be applied as part of user instruction. But, as the guidelines have the objective to be applied not only to body, but also to objects instructions, we implemented prototypical tools related to objects too.

Most AR approaches focus on part of what it was presented in this chapter: focus on visualization or instruction or a specific improvement in AR development. We presented that it was possible to put these tools together to help the AR designer to identify interesting techniques that can be used in their approach.

5 EVALUATION STAGE

This stage aims to collect feedback from users and to identify how useful instructions were in helping users to accomplish defined tasks. Possible evaluation methods include subjective user feedback through interviews and surveys, and user studies. In our proposal, which brings yet another contribution to the AR area, due to the lack of works related to body instructions, the test scenarios will focus on test cases related to body movements. We defined a pilot test to validate this proposal; first we will discuss how physiotherapists and yoga teachers traditionally apply instructions to understand how these professionals work. The second kind of evaluation was done through the analysis of three different types of AR applications to understand how instructions are worked and if the guidelines could be useful and applied on them.

5.1 Understanding How Instructions are usually applied to Body Movements Activities

To apply our AR instruction system, first, we needed to understand how instructions are applied by yoga teachers and physiotherapists (we call them 'instructors' from now on). We tried to figure out their limitation, difficulty and feedback that they get from students/patients. A questionnaire with sixteen questions was applied with fifteen (15) Brazilian instructors; the questions are presented in Table 6. In the left column, questions are presented and in the right column, the possible answer, if it is not an open question (in this case, the instructor must write a text as a way to answer the question).

Table 6. Questions applied to fifteen Brazilian instructors

Question	Possible Answers
1. Which is your main activity?	• Yoga teacher • Pilates instructor • Physiotherapist • Physical educator
2. How long do you work in your area?	• Less than 5 years • Between 5 and 10 years • More than 10 years
3. How many patients/students do you attend each week?	• Less than 10 • Between 11 and 30 • Between 31 and 50 • More than 50
4. During the classes/sessions, how	• By demonstration

Question	Possible Answers
patients/students understand a new movement?	• By verbal instructions • By images • Other: ______________ (*to specify*)
5. In your activity, is it important to have activities outside the classes/sessions (activities that the students/patients can do by themselves)?	• Yes • No • Maybe
6. If you answered 'No' or 'Maybe' in question 5, specify the reason.	
7. Are you used to give activities to the students/patients to do at home?	• Yes • No
8. If you answered 'No' in question 7, justify.	
9. If you answered 'Yes' in question 7, how do you give the instruction to them?	• By images/text • By demonstration
10. If you answered 'by images/text' in question 9, how do you give the perception of depth, velocity and direction?	
11. If you answered 'by images/text' in question 9, what kind of resources do you apply more?	• Text • Images • Draws • Other: ______________(*to specify*)
12. What kind of activities are most interesting to students/patients to do by themselves?	
13. How do you evaluate if the student/patient did correctly the movement?	
14. How is the feedback of the students/patients about the activities done by themselves?	• Satisfied • Unsatisfied
15. Describe more, if you answered 'Unsatisfied' in question 14.	
16. What are the most important issues during the activities done by the students/patients?	• Number of repetitions • Different sequences of movements • Other: ______________(*to specify*)

Questions 1, 2 and 3 are related to the background of the instructor and allow us to identify if the instructions are managed in a different way, in agreement with the activity performed, the experience the instructor has in his/her field and the number of students/patients he/she attends per week. 1 yoga teacher with more than 10 years of experience and more than 50 students, 1 Pilates instructor with less than 5 years of experience and a number of patients between 11 and 30, 2 physical educators with more than 10 years of experience and at least 30 students, 1 physical educator with less than 5 years and the number of students between 11 and 30; 8 physiotherapists with less than 5 years of activity and most of them with less than 30 patients; 2 physiotherapists with more than 5 years and less than 50 patients. A summary of this information is presented in Table 7.

Table 7. Summary from Brazilian instructors' answers

Question	Answer
Main activity?	• 6.7% Yoga teacher • 6.7% Pilates instructor • 66% Physiotherapist • 20% Physical educator
2. How long do you work in your area?	• 60% Less than 5 years • 20% Between 5 and 10 years • 20% More than 10 years
3. How many patients/students do you attend each week?	• 40% Less than 10 • 26.7% Between 11 and 30 • 26.7% Between 31 and 50 • 6.7% More than 50

All instructors surveyed agreed that it is important that the patients/students do exercises by themselves and/or at home; 13 professionals give the instructions to do at home, usually by demonstrations (only 2 by text/images) – see Fig. 62.

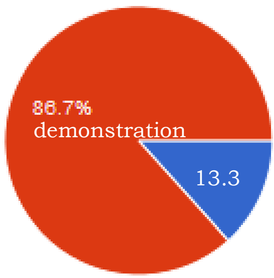

Fig. 62. Summary of ways to present new instructions

The evaluation and feedback of the movements are done mainly verbally and through demonstrations. Based this understanding of how instructions are usually applied to body movements activities, our proposal has the goal to avoid wrong movements and motivate the users by the application of our AR instructions system.

5.2 IKAPP LIBRARY

IKapp is a rehabilitation support system developed by GAMA et al. [2012a, 2012b] composed by a body tracking and a biomechanical movement analysis modules. The first module is responsible for the extraction of the body skeleton data using a RGB-D sensor and the recognition of body movements. The second module is responsible for the analysis of the user's movements from the perspective of biomechanical parameters including planes and angles. To give a feedback to the users the IKapp library provides a game where the user movement controls a dolphin. The game interface can be seen in Fig. 63: the user's arm movement controls the red dolphin shown in the screen. IKapp allows the physiotherapist to register a therapy protocol in the system, and only the exercises registered in the system allow the user to control the game. Beyond, the system recognizes if the movement performed by the user is correct according to the registered movement.

Fig. 63. Game interface of the IKapp library

Games are an interesting way to motivate the users. But they do not correct and teach the users the correct movement. So, instead of applying the game interface from iKapp, we applied our instruction module; we applied a mirror view: the user sees his/her image overlapped with the instructions.

IKapp implements the movement recognition through the use of checkpoints – a set of poses that compose the full movement. The library informs the actual state of the movement in agreement with the complete movement, and our AR instruction system can inform correctly the users the best way to reach the next pose or to visualize the full movement.

The mirror view can be implemented with a TV or a projector. We did four tests with a TV and 10 tests with a projector. Fig. 64 illustrates the setup environment to the pilot test with a projector. The RGB-D sensor was put in front of the user; a computer was used to process the data.

Fig. 64. Setup environment to the pilot test

5.3 PILOT TEST

As the instructors reported to use printed instructions (text/images) and demonstrations to guide their students/patients, we will compare our system with these kinds of instructions. The testers will answer a questionnaire related to the instructions in the end of the session.

As there are different kinds of instructions identified in our work, and we can have different combinations of them, we selected the most representative ones to validate the guidelines. The instructions applied were:

- *Text/2D draws*
- *Arrows*
- *Edges*
- *Avatar*
- *Ghosting*

Bartlett [2007] illustrates and describes different types of body movements. We selected two movements to be used in our pilot test: shoulder abduction and hip flexion. These movements (or exercises) are commonly applied in Physiotherapy, Yoga and workouts. They are related with different body parts and require a different movement from the user. Besides, these movements can be captured by the IKapp library.

Shoulder abduction [LABORATORY ERGONOMICS, 2015] is a lateral movement away from the midline of the body, moving the upper arm up to the side away from the body (Fig. 65, left). Hip flexion is the movement that puts the upper leg upward to the front (Fig. 65, right) of the body.

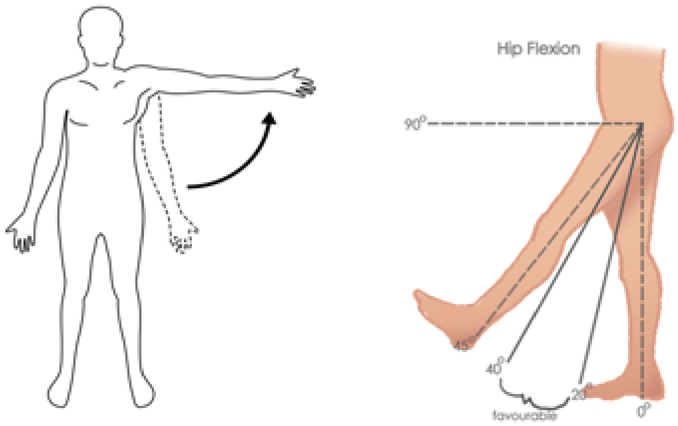

Fig. 65. Illustration of shoulder abduction (left) and hip flexion (right)

Despite seeming easy to do, These movements, may cause injuries if done wrongly, as reported by the *National Institute of Arthritis and Musculoskeletal and Skin Disease* (NIH)[2]: dislocation, separation, rotator cuff disease, rotator cuff tear. Besides, after having an accident or aging problems, people can have difficulty to do these movements.

The entire system for movement recognition and instructions presentation was running on a computer with an i5 processor and 4GB of memory. The pilot test was applied to fourteen users.

5.3.1 PROTOCOL

For each user participating in the pilot test firstly a demonstration of the shoulder abduction movement was performed. Then they were asked to repeat the movement for each instruction evaluated. The sequence applied is illustrated in Fig. 66: text/2D draws (left image on top), arrows (middle image on top), edges (right image on top), avatar (left image on bottom) and ghosting (right image on bottom).

[2] http://www.niams.nih.gov/Health_Info/Shoulder_Problems/shoulder_problems_ff.asp

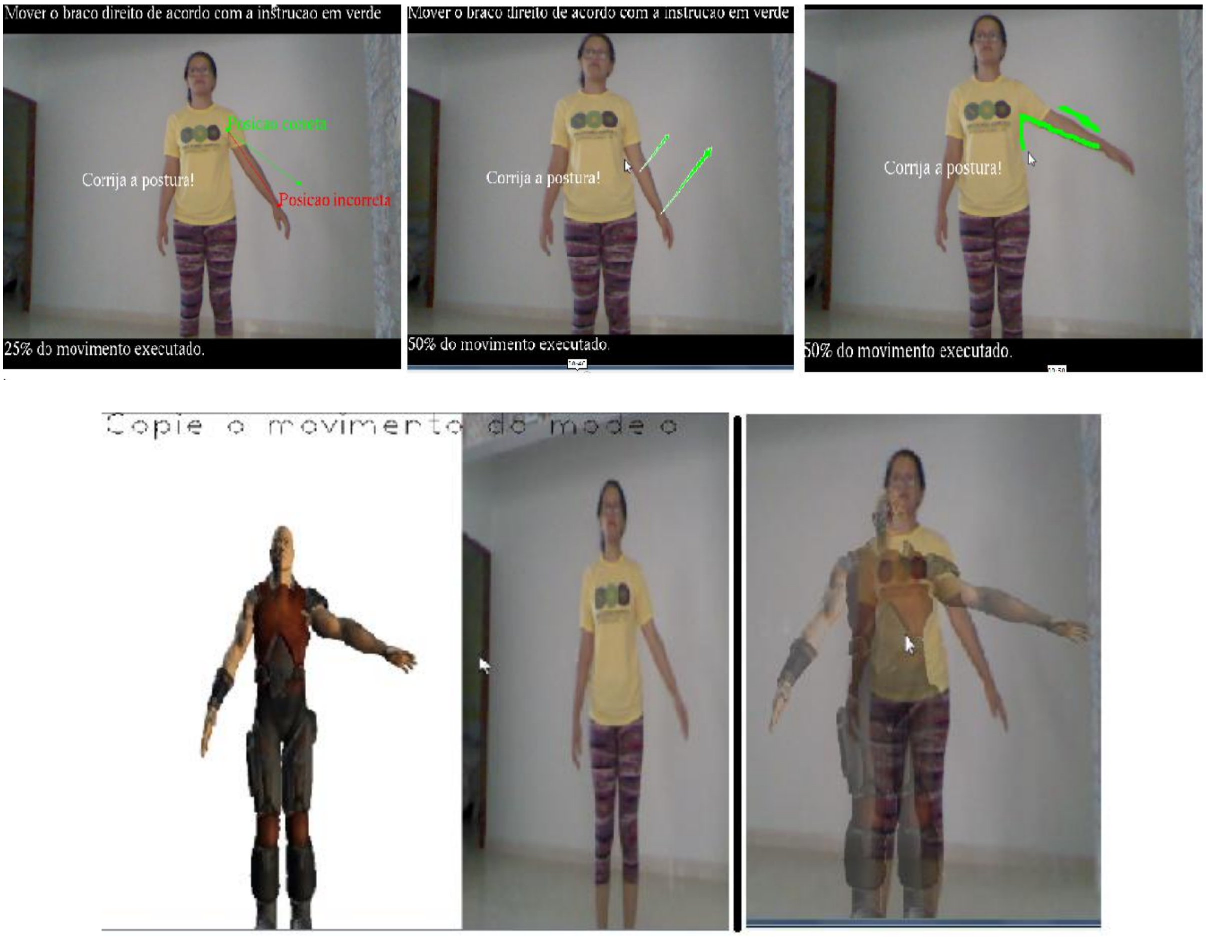

Fig. 66. Sequence of instructions applied in the pilot test: text/2D draws intructions, arrow intruction, edge intruction, avatar instruction and ghosting intruction

As the full movement is composed by a set of poses (checkpoints), it was explained to the users that the system gives an illustration for the next pose, until the user reaches the final pose. An example of sequence of text/2D instruction is illustrated in Fig. 67: the correct arm position is illustrated in green, the woman must put her arm (red lines) in the right position (left image); when she reaches the green line, the system presents the next pose (middle image). To help her, a bottom message presents the status of the actual pose in agreement with all the poses that belongs to the movement. In the end, the system gives a message that the exercise was done (right image).

The system only allows the user to go to the next instruction, when the user does the full movement (shoulder abduction movement) or after a predefined time (15 seconds). Some messages can be presented to help the user, for example 'correct your posture'.

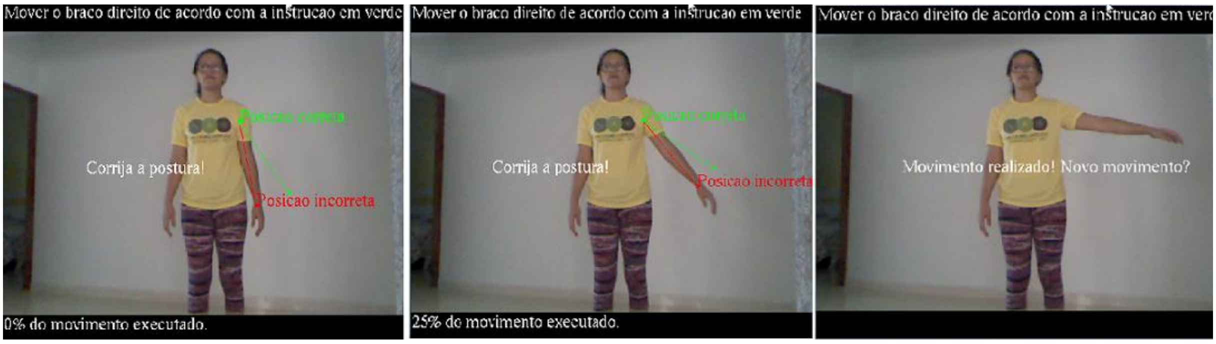

Fig. 67. Example of sequence of text/2D draw instructions related to shoulder abduction movement

The hip flexion movement was demonstrated to user and the same procedure was done to this movement.

Despite the fact that the system can accept the repetition of movements for each instruction, the user does only one movement per instruction; as we had five instructions in the pilot test, the users do at least five repetitions from the same movement. As there are two movements, each user does 10 movements. The pilot test aimed at: fast user evaluation, to have a initial feedback and improvements to do before doing the specific tests scenarios (see the future works section). In the end of the session, the user answered a questionnaire with 24 questions - objective and subjective ones. The objective answers are based on a 1-5 Likert scale questionnaire.

Table 8 presents the questions applied.

Table 8. Questions applied with the testers in the end of the session

Question
Personal information (age, gender, physical limitation, level of studies, level of informatics knowledge, level of electronic games knowledge)
A. About the messages and 2D draws 1. They were useful to give tips and/or use the system □ strongly agree □ agree □ neutral □ disagree □ strongly disagree 2. The color was adequate □ strongly agree □ agree □ neutral □ disagree □ strongly disagree 3. I did not have difficulty to read □ strongly agree □ agree □ neutral □ disagree □ strongly disagree 4. They were presented on the right position □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
B. About the arrows

Question
1. They helped to identify the movement direction □ strongly agree □ agree □ neutral □ disagree □ strongly disagree 2. The color was adequate □ strongly agree □ agree □ neutral □ disagree □ strongly disagree 3. I did not have difficulty to read □ strongly agree □ agree □ neutral □ disagree □ strongly disagree 4. They were presented on the right position □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
C. About the avatar that illustrates the next position to reach 1. Its appearance did not influence the movement execution □ strongly agree □ agree □ neutral □ disagree □ strongly disagree 2. If its appearance was other or If I can change it, it would be a motivation and lead to a better comprehension of the movement □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
D. The emphasis/highlight of the body part to move helped to identify correctly the body part to move □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
E. The avatar helped to indicate the correct direction □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
F. The avatar was the best way to have a right view of the movement □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
G. The ghosting avatar was better than the avatar because it keeps the scene context □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
H. The ghosting avatar was better than avatar because it presented better the movement □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
I. Would you use this kind of system to perform body exercises (workouts, physiotherapy, yoga, dance)? □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
J. Do you think that the presented instructions were in agreement with the presented movements? □ strongly agree □ agree □ neutral □ disagree □ strongly disagree
K. Did you have some difficulty?
L. Did you think that the system ran fast?
M. Did you have some suggestion or critics about this kind of system?

To define the users sample size, as well as to analyze the consistency and reliability of the answers, the Cronbach's alpha coefficient [CRONBACH, 1951] was applied. It measures the correlation between the given answers to verify, as defined in statistics area, the internal consistency. The coefficient varies between 0 and 1; the minimum good value is 0.7, meaning that the result is acceptable and reliable.

The Cronbach's alpha α is defined by:

$$\alpha = \left(\frac{k}{k-1}\right) \times \left(1 - \frac{\sum_{i=1}^{k} s_i^2}{s_t}\right)$$

where:

- k is the number of the items of the questionnaire;
- s_i^2 is the variance of each item;
- s_t is the total variance – sum of all variances.

the website StatsToDo[3] was used to calculate the sample size. The needed parameters are:

- *Type I error* (significance level) – 0.05 means less than 5% of error;
- *Power* - is the probability of rejecting the null hypothesis when the alternative hypothesis is true [MCCRUM-GARDNER, 2010] – the minimum value is 0.8 or 80%;
- *Number of items* – it is 17 (17 questions to analyze);
- *Expected Cronbach's Alpha* – 0.7.

The sample size estimation for our pilot test was 14. More details about sample size estimation can be obtained in MCCRUM-GARDNER [2010] and COHEN [2011].

5.3.2 RESULTS AND DISCUSSION

The Cronbach's alpha coefficient for our questionnaire was *0.7133* reaching the minimum good coefficient required.

The AR instruction system was applied to 14 people – 9 men and 5 women – with ages between 17 and 48, and different levels of education and professions (physiotherapist, students, teachers, and economist).

[3] Available at https://statstodo.com/SSiz1Alpha_Pgm.php#Single%20calculation%20:%20sample%20size%20estimation Last view: January, 2016.

Table 9 presents the mean and standard deviation values for all answers of the questionnaire.

All users answered that they agree or strongly agree that they would like to use this kind of system to do body movements (question I). They also agree that the instructions are adequate for this kind of task (question J).

For *indication of movement*, we proposed the use of arrows. Question B1 asked if they are fulfilling this goal, and users agreed. The color presented was fine (question B2) as well as arrows position (question B4). But some users reported that the arrows head must be bigger to identify them better; this problem is related to question (B3).

We asked whether the avatar could be used to indicate the direction of movement, as the user has the option to follow it (question E). In spite of being a 3D representation of the movement, we thought that the answer will be better than a 2D representation, but the answers were worse. In the same way the avatar ghosting showed lower results than the use of arrows (question G). About 21% of users did not agree that the avatar is the best way to visualize the movement and 28% were indifferent (question F). We analyzed the answers presented in questions C1 and C2 to identify a path to follow; in both questions, we had extreme answers, but some users that answered 1 or 2 in question C1 had the tendency to put a higher score in question C2, so a better avatar representation must be tested in further studies. Besides, the avatar instruction does not correct them as the arrows and messages sent.

The use of an avatar was proposed to validate the guideline *management of depth* related to body movements; avatar ghosting was proposed to the guideline *management of occlusion* with body movements. 71% of the users agreed that the avatar ghosting is better than avatar (question H). Through avatar ghosting, the user can easily move the body part related to the movement in a right position in agreement with the overlapped avatar; besides the context is kept.

The guideline *emphasize parts* to focus the user attention on the right body part was validated in question D; about 78% agreed that this was a nice way to show the instruction.

The guideline *feedback* could be validated by the messages sent by the system. Question A1 asked directly if the messages were useful and helped users during the exercises. All users set equal or more than 4 (agree) to this sentence.

About the movement recognition, the system recognizes easier the shoulder abduction than the hip flexion. As we did changed some modules of the IKapp library, this could have affected the recognition of frontal movements; this is a limitation of the Kinect SDK too. We did not realize bad evaluation from users, when at least the shoulder abduction movement was recognized. But the users give less score compare to users that have both movements recognized. In spite of this, it was the evaluation was good.

Table 9. Mean and standard deviation of the questionnaire answers for the pilot test

Question	Mean	Standard Deviation
A. About the messages and 2D draws		
1. They were useful to give tips and/or use the system	4.5	±0.52
2. The color was adequate	4	±1.1
3. I did not have difficulty to read	3.86	±1.17
4. They were presented on the right position	4.36	±0.87
B. About the arrows		
1. They helped to identify the movement direction	4.36	±0.84
2. The color was adequate	4.36	±0.74
3. I did not have difficulty to read	3.93	±1.00
4. They were presented on the right position	4.21	±0.58
C. About the avatar that illustrates the next position to reach		
1. Its appearance did not influence the movement execution	3.50	±1.70
2. If its appearance was other or If I can change it, it would be a motivation and lead to a better comprehension of the movement	3.71	±1.38
D. The emphasis/highlight of the body part to move helped to identify correctly the body part to move	4.14	±1.17
E. The avatar helped to indicate the correct direction	3.93	±1.00
F. The avatar was the best way to have a right view of the movement	3.50	±1.09
G. The ghosting avatar was better than the avatar because it keeps the scene context	3.93	±1.00
H. The ghosting avatar was better than avatar because it presented better the movement	3.86	±1.03
I. Would you use this kind of system to perform body exercises (workouts, physiotherapy, yoga, dance)?	4.71	±0.47
J. Do you think that the presented instructions were in agreement with the presented movements?	4.21	±0.43

5.4 AR DESIGNERS ANALYSIS

Our final validation was done with developers from three different types of AR applications. Thus, we could analyze if our guidelines could be useful to them to apply instructions and improve the user experience in their application. All the three applications were created in VoxarLabs[4] – a research group from the Federal University of Pernambuco that has AR subject as topic of research and development. Each AR application has at least five people involved with different roles (some of them can assume more than one role), so we asked a set of questions to the main developer to evaluate the instruction process. The three applications were: mirrARbilitation, Bare Hand Natural Interaction with Augmented Objects and ARBlocks.

5.4.1 OVERVIEW OF AR APPLICATIONS ANALYZED

The mirrARbilitation[5] is an application to motivate and help the users to do physiotherapy exercises. The users has the view of themselves on a mirror and the instructions are done by moving 2D objects (see Fig. 68 a man must move a ball to put it in a basket). Besides using 2D objects, texts are also applied to help the users to understand the movement. The user skeleton data is obtained from the application of a depth sensor and they apply the iKapp library (explained in previous chapter) to check the pose.

Fig. 68. Example of a session from mirrARbilitation

The second application analyzed was related to interactions between

[4] http://cin.ufpe.br/~voxarlabs/Home.html. Last view: April, 2016.

[5] https://www.youtube.com/watch?v=nvNqkJ0l2yI. Last view: April, 2016.

augmented objects on real scenes with a real hand; the title of the application is Bare Hand Natural Interaction with Augmented Objects [FIGUEIREDO et al., 2013]. An example of task in this application is presented in Fig. 69: the user must move the red cub: left image, the user moves the red cub; center image, the initial state; and, right, the final position. The application considers the information of a color image and a depth map to assign 3D positions for the hands and the fingers used for gestures. To facilitate the presentation of the discussion on the next session, we will call this application HandNIAR.

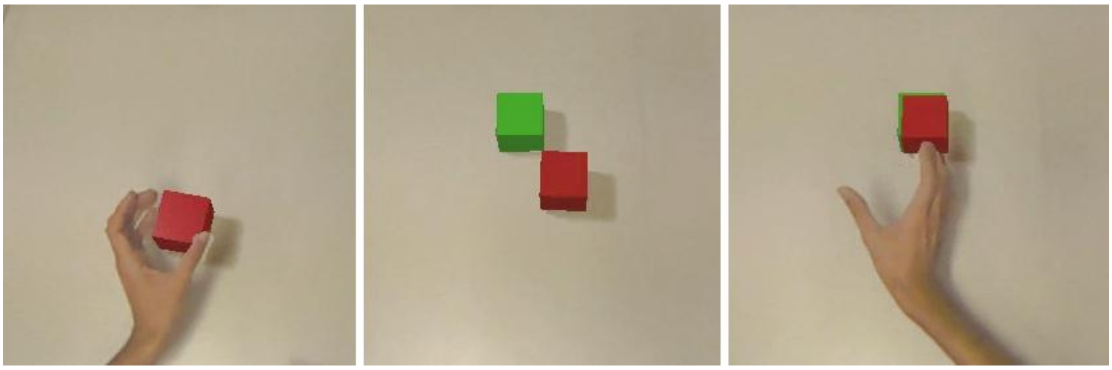

Fig. 69. Bare hand natural interaction with augmented objects

The third application is ARBlocks [Roberto et al., 2013] which is based on projective augmented reality and tangible user interfaces aiming early childhood educational activities development. The instructions are presented on real blocks and different kinds of tasks can be applied on it, but it was developed a simple game aiming to help teachers in early childhood literacy; in this game, drawings (animals or objects) are exhibited in half of the blocks, while their respective words appear in the others. The objective of this game is to place together the blocks with word and the drawing. When the children make the correct association, a green rectangle is presented to them – see Fig. 70.

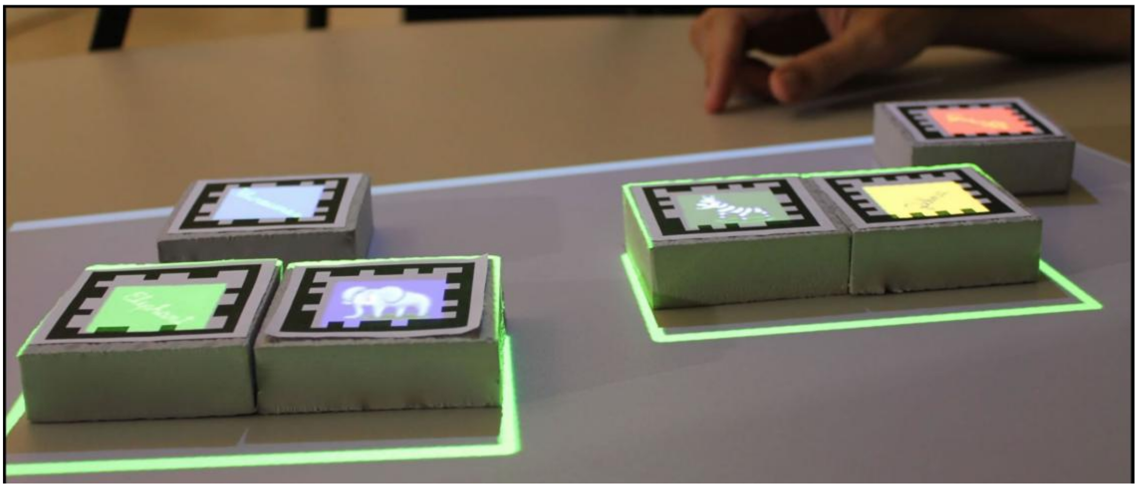

Fig. 70. A game to children built on ARBlocks

5.4.2 *PROTOCOL*

To each application, we asked the 28 questions (objective and subjective) to understand and to evaluate if the guidelines could be useful in the development of it. These questions are presented in Table 10.

Table 10. Questions applied to AR developers

Question
1. What is the name/title of your AR project/application?
2. Is there a link about it?
3. Are there publications about it? If your answer was yes, please inform the reference of the most representative.
4. How many people worked in the project/application?
5. What is the main subject/topic of your application?
6. How long have you been working in the project/application?
7. Did you do user evaluation?
8. Could you write some positive aspects given by users?
9. Could you write some negative aspects given by users?
10. What kind of instructions is used in your project/application?
11. How did you choose or define the instructions?
12. Any of your instruction has the goal to give the idea of movement or direction to the user?
13. If you answered 'yes' in the previous question. Please inform the instructions that have this goal in your application.
14. Any of your instruction has the goal to highlight or emphasizing part of the scene to help the user?
15. If you answered 'yes' in the previous question. Please inform the

Question
instructions that have this goal in your application
16. Any of your instruction has the goal to manage depth or occlusion?
17. If you answered 'yes' in the previous question. Please inform the instructions that have this goal in your application
18. If had a set of guidelines related to the use and application of instructions in the development of AR applications, would you read it with the goal to apply it in your application?
19. Read the text: "indication of movement - it is important that your application indicate the correct path or the correctness of the movement; you could apply, for example, 2D/3D arrows and/or animation". Have you already applied this information on you application?
20. Read the text: "indication of movement - it is important that your application indicate the correct path or the correctness of the movement; you could apply, for example, 2D/3D arrows and/or animation". If you had this information before coding your application, would this information help you to apply this kind of instruction?
21. Read the text: "emphasize parts - it is important that your application indicate/emphasize parts of an object or a body to be moved or changed or get the user attention; you could apply edges emphisaziment, animation and/or diminished reality". Have you already applied this information on you application?
22. Read the text: "emphasize parts - it is important that your application indicate/emphasize parts of an object or a body to be moved or changed or get the user attention; you could apply edges emphisizement, animation and/or diminished reality". If you had this information before coding your application, would this information help you to apply this kind of instruction?
23. Read the text: "allow different kinds of visual appearance attributes - the visual appearance of the AR instructions should agree with the environment conditions (light, pattern of colors and so on). The user should remain in control over the appearance, because some tasks could have specific patterns to follow; you could apply, for example, color changes". Have you already applied this information on you application?
24. Read the text: "allow different kinds of visual appearance attributes - the visual appearance of the AR instructions should agree with the environment conditions (light, pattern of colors and so on). The user should remain in control over the appearance, because some tasks could have specific patterns to follow; you could apply, for example, color changes". If you had this information before coding your application, would this information help you to apply this kind of instruction?
25. Read the text: "feedback - the process of giving an instruction must deal with feedback for users to convey if the user is proceeding

Question
correctly, what must be changed, or if there are alternatives. It is essential that such suggestions be presented in real-time; for example, you could apply text, 2D/3D draws". Have you already applied this information on you application?
26. Read the text: "feedback - the process of giving an instruction must deal with feedback for users to convey if the user is proceeding correctly, what must be changed, or if there are alternatives. It is essential that such suggestions be presented in real-time; for example, you could apply text, 2D/3D draws". If you had this information before coding your application, would this information help you to apply this kind of instruction?
27. Read the text: "Management of occlusion and depth - Instructions conveying three-dimensional information, such as a disassembly task, must incorporate visual cues to let the user understand occlusion, depth relationships and distances in general; you could apply, for example, avatar, ghosting technique and/or x-ray vision". Have you already applied this information on you application?
28. Read the text: "Management of occlusion and depth - Instructions conveying three-dimensional information, such as a disassembly task, must incorporate visual cues to let the user understand occlusion, depth relationships and distances in general; you could apply, for example, avatar, ghosting technique and/or x-ray vision". If you had this information before coding your application, would this information help you to apply this kind of instruction?

5.4.3 *RESULTS*

The questions 1 to 9 were meant to obtain basic knowledge about the application, such as the main subject managed, number of people working in it, time of developing. Some aspects caught our attention, such as the time of development; it is necessary at least one year of working (HandNIAR); mirrARbiliation has been developed for five years and ARBlocks for two years. There are between five to seven people involved in each application. A summary of these answers are presented in Table 11.

Table 11. Summary of basic information from AR applications

Application	People involved	Main topic	Time of working
mirrARbilitation	5	Physiotherapy	5 years
HandNIAR	7	Education, Games	1 year
ARBlocks	5	Education	2 years

All of them did user evaluation (question 7), and the negative issues related by users (question 9) in mirrARbilitation and HandNIAR are important aspects that the guidelines try to manage. In mirrARbilitation some users said that they have difficulties to understand the movement, and in HandNIAR they have experienced problems to realize better the depth perception.

A summary of the instructions applied in each application is presented in Table 12 (question 10). In spite of being a common type of instruction, texts are not applied to HandNIAR, but they get the depth information of the scene to apply shadows to give to users the depth information. ARBlocks apply text, images and image processing techniques to highlight or get the user attention to the task.

Table 12. Summary of instructions applied to AR applications

Instruction	Application
Text	mirrARbilitation, ARBlocks
Images	mirrARbilitation, ARBlocks
Image processing techniques	ARBlocks, HandNIAR
2D virtual representation of the real objects	mirrARbilitation
3D virtual representation of the real objects	HandNIAR
Other (shadows to give depth perception)	HandNIAR

In question 1 we asked how they defined these instructions: mirrARbilitation defined them through discussion with developers and physiotherapists, HandNIAR defined them by the team experience on these instructions and ARBlocks defined by a combination of team experience on these instructions and discussions with developers and experts in the field that AR application will be used. Each application followed its own path; there is no agreement in the way to define the kinds of instruction to apply. All developers agree that if there was a set of guidelines or indication of ways to apply instructions, they would read it to try to implement in their application (Question 18).

From question 12 to 17 we asked if and how the applications manage movement, depth, feedback, emphasize and occlusion – the answers could help us to validate if the examples of instructions that we proposed for each guideline is in agreement with their choice. mirrARbilitation and ARBlocks applies text as a tool for give the idea of feedback about what the user needs to do.

mirrARbilitation manages only movement information and the feedback is given by text. To give the idea of movement to users, it applies images and 2D objects. As users, in their evaluation, related problems to perceive correct movement, we believe that the 2D objects implemented were not enough; the application of 2D or 3D arrows proposed in the guidelines would be useful. Table 13 summarizes the relation between the instructions proposed in the guidelines and the instructions implemented in mirrARbilitation.

Table 13. Comparison between the guidelines and the instructions applied in mirrARbilitation

Guideline	Examples of Instructions	mirrARbilitation
Indicate movement	Arrows, stroboscopic motion, key poses, images, animation	Images
Emphasize parts of an object or a body to be moved or changed or get the user attention	Edge outlines, diminished reality, animation, color harmony	Not managed
Allow different kinds of visual appearance attributes	Color changes, 2D/3D objects, color harmony, animation	Not managed
Feedback	Text, color changes, arrows	Text
Management of occlusion and depth	Avatar, ghosting techniques/x-ray vision, key poses, arrows	Not managed

HandNIAR does not apply any instruction with the idea of movement. But it applies image processing techniques to manage depth, occlusion and emphasis of areas. The feedback is an issue that must be improved. Table 14 summarizes the relation between the instructions proposed in the guidelines and the instructions implemented in the application.

Table 14. Comparison between the guidelines and the instructions applied in HandNIAR

Guideline	Examples of Instructions	HandNIAR

Indicate movement	Arrows, stroboscopic motion, key poses, images, animation	Not managed
Emphasize parts of an object or a body to be moved or changed or get the user attention	Edge outlines, diminished reality, animation, color harmony	Image processing techniques
Allow different kinds of visual appearance attributes	Color changes, 2D/3D objects, color harmony, animation	Not manage
Feedback	Text, color changes, arrows	Not manage
Management of occlusion and depth	Avatar, ghosting techniques/x-ray vision, key poses, arrows	Image processing techniques

ARBlocks manages movement and emphasis. Texts, images and animation are the types of instructions applied by it. Images and animations were applied to give the idea of movement and emphasis, but the text instruction is better placed as a feedback or complementary instruction, as proposed by the guidelines. Table 15 summarizes the relation between the instructions proposed in the guidelines and the instructions implemented in the application.

Table 15. Comparison between the guidelines and the instructions applied in ARBlocks

Guideline	Examples of Instructions	ARBlocks
Indicate movement	Arrows, stroboscopic motion, key poses, images, animation	Text, images, animation
Emphasize parts of an object or a body to be moved or changed or get the user attention	Edge outlines, diminished reality, animation, color harmony	Image processing techniques
Allow different kinds of visual appearance attributes	Color changes, 2D/3D objects, color harmony, animation	Animation
Feedback	Text, color changes, arrows	Text, animation
Management of occlusion and depth	Avatar, ghosting techniques/x-ray vision, key poses, arrows	Not manage

Questions 19 to 28 present the guidelines: there is a description of them and possible types of instructions that could be applied. To questions 20, 22 and 26, that present the guidelines to manage movement, feedback and emphasis, all the developers agree that if they had these data before, the information presented that would help them in the development. ARBlocks and mirrARbilitation agreed that the information presented in question 24, to allow different kinds of visual appearance attributes, would help them. Only HandNIAR the information about the management of depth and occlusion would help it (question 28), but this application is the only one that manages depth and occlusion.

5.4.4 *DISCUSSION*

The topic of each application is different, and the way of giving instructions too, we could have important analysis and information to validate and improve the guidelines. Two applications interact with virtual objects (HandNIAR and mirrARbilitation), one with real objects (ARBlocks). mirrARbilitation is the only one which is important that the user sees himself/herself doing the exercises to receive the feedback; in HandNIAR is necessary only his/her hand to get the positions to interact with virtual objects. Even though there were three different AR applications, the developers answered all the questions and they could relate their applications with the guidelines. As our initial goal was to define a set of guidelines to be independent of context seems that we are in right way.

mirrARbilitation has less fewer types instructions in comparison to the others, because the application focuses on the movement recognition; but the way that the application instruct can cause bad user evaluations (as presented in previous session). But the application can improve significantly the instruction process by the application of the other guidelines (only two are implemented).

HandNIAR is the only application that manages depth and occlusion. It gets the depth map, applies photorealistic rendering to create shadows and management of the occlusion. Our guidelines propose different techniques, but after the analysis of them, we can put photorealistic rendering as more one technique that can be applied to manage depth and occlusion. But, we need to remember that some users complained about depth perception, so a combination of our initial proposal techniques and photorealistic rendering can help the instruction process.

ARBlocks illustrates for each guideline used, a technique that was proposed in the respective guideline. And the users' complaints were not related to instructions, but to other techniques issues, like light conditions. In this way, we believe that this is more one confirmation that we defined right representative techniques for the guidelines.

5.5 Conclusion

This chapter had the goal to present two different aspects that can be used to evaluate the guidelines: the end user evaluation and the AR designer/developer evaluation. For the first evaluation, we created an AR system that implemented a representative instruction for each guideline; in this way, we verified the end user acceptance, understanding of the task and/or movement in agreement with the instruction implemented. In spite of having good scores, some instructions must be improved, as avatar and avatar ghosting. These instructions give us 3D view of the movement, but the appearance of them caused ambiguity.

In the second evaluation, the guidelines were validated by the designer's view. As we created an AR application based on the guidelines, and this system was specific to body movements, we saw the necessity to validate the guidelines to applications that use body and movements, as our objective is to have guidelines to be applied in both cases. So, this second evaluation, for three different AR applications (one related to body movements, two related with objects interaction), the developers presented the instructions applied, and we identified if and how the instructions and guidelines proposed were implemented. Besides, the designers/developers informed the need of some material to help them to the application of instructions. And finally, the guidelines seem robust to be applied to body and objects.

6 CONCLUSION

This thesis investigated whether it was possible to define general guidelines for the development of instructions in AR approaches; through the analysis of visualization and instructions applied to AR, we proposed five guidelines that can be used as guidance. Of course, that in agreement with the objective of the application, there is no necessity to follow all the guidelines, but part of them; for example, there are applications/tasks that are not important emphasize parts but indicate direction and others the, opposite.

From our analysis of visualization techniques we identified that there is a large number of visualization techniques applied to AR, but their use for presenting instructions may be explored further. The approaches focus on the development or adaptation to AR, but few focus on how to put them in practice. We suggested how some of them can be used; for example, we identified how Diminished Reality and Color Harmony can be applied as part of AR instructions.

We identified that some visualization techniques can be applied with specific objectives: Ghosting is a very interesting technique to keep the user context, but it is should be applied in combination with other instructions (such as arrows and texts) to help the user understand what to do. Similarly, edges are better applied in combination with arrows.

We identified different kind of instructions applied in AR applications. Each application has its own way to help the users during a task. The most common kind of instruction applied to AR is text, but it works fine usually as complement instruction. Its combination with animation, images and virtual objects seems to be more effective.

In this thesis we have presented many works related to objects instructions and few related to body instructions. The majority of approaches related to body instructions were developed in the last few years; we believe that this is due to the decrease of the price of equipments We developed prototypes for objects and body instructions; we have also created a specific system and did a user evaluation for body movements, due to he lack of related works. In order to evaluate whether the guidelines could be applied in different scenarios, we have evaluated three distinct AR applications. From their analysis , our tests, results and prototypes created, we concluded that the guidelines could be applied in both.

It is important remark that we have proposed five guidelines to help in the development and instruction application to improve the user experience, but some of the guidelines and instructions cannot be applied to all cases – for some applications emphasis on parts are not relevant, but the indication of movement or depth information are more important. But, at least, the developer has some material to help him/her in the development.

6.1 CONTRIBUTIONS

The main contributions of the work presented in this thesis are:

- An AR instructions system with a set of instructions that can be used for body instructions;
- Definition of a set of visualization techniques that can be applied to body and object instructions;
- A set of guidelines to apply in AR instructions;
- A proposal to apply visualization techniques, such as Diminished Reality, Color Harmony and Ghosting, as part of AR instructions;
- Publications directly related to this work:
    - ROLIM, C. and TEICHRIEB, V. 2012. A viewpoint about diminished reality: is it possible remove objects in real-time from scenes? 2012 14th Symposium on Virtual and Augmented Reality. (May. 2012), 141–146.
    - PADILHA, A., ROLIM, C. and TEICHRIEB, V., 2013. The ghosting technique applied to augmented reality visualization. In XV Symposium on Virtual and Augmented Reality (SVR). pp. 159–166.
    - ROLIM, C., SCHMALSTIEG, D., KALKOFEN, D. and TEICHRIEB, V., 2015. [POSTER] Design Guidelines for Generating Augmented Reality. In IEEE International Symposium on Mixed and Augmented Reality. Fukuoaka, Japan.

6.2 FUTURE WORK

Our guidelines were defined based on the analysis of object and body instructions, the AR approaches manages one specific issue per time (body or object), an evaluation of the guidelines when both are involved in the task is necessary. Besides, tasks that are done in collaboration – many users involved – could have specific issues that must be managed.

The study of cognitive factors when the guidelines are applied is interesting to check the influence of the task, environment and previous user knowledge. The approach proposed in this thesis applies only one RGB-D sensor to get depth and color information of the scene/user. It would be interesting to evaluate the application of more than one RGB-D sensor to give to the user different viewpoints of the task or illustrated procedure.

As presented in this thesis, FURMANSKI et al. [2002] define a set of guidelines to manage the depth perception in x-ray visualization; as a future work, it would be interesting to do an evaluation of the ghosting and x-ray visualization applied to AR instruction in agreement with their principles.

In some contexts, the system does not provide a perfect visualization (e.g. an arrow could be presented in different directions). In this context, it would be interesting to work on an editor do help the user to edit the visualization sequence. It is necessary to improve the visualization and instructions with objects.

The avatar applied here is a male model (dressed like a soldier); we have the goal to improve it, so that the user can choose a model based on gender, age, kind of body and clothes. Another possibility is the use of a 3D avatar model of the own user (a 3D reconstruction of him/her).

Some improvements on the visualization techniques could be implemented to help different kinds of contexts, for example, improvements to indicate the velocity and level of danger of the movement. We have the intention to look for activities or movements done in collaboration with others users (exercises done by groups of users) to verify what kind of improvements would be necessary.

A pilot test was carried out witha general public. An investigation to create specific user studies to physiotherapy, dance, yoga, work outs could be interesting.

We have tested one instruction at a time, but they can be used in combination. Another user study to verify the combination of the instructions must be done.

The pilot test was applied to users with ages varying between 17 to 48, but there is a significant part of the population that is more than 60 years and has difficulties going out or doing exercies or physiotherapy sessions. Besides, this population may have vision and body limitations. An investigation about the better kind of instructions and how they must be presented would be interesting.

We could not evaluate occlusion, but this is an important issue specially to object instructions. An evaluation different ways to present instructions as well as their behavior with occlusion must be done.

We will integrate our developed Diminished Reality, Color Harmony and 3D Reconstruction approaches to the body and object instructions system. Besides, it is interesting to create user studies with these techniques as a way to give instructions. Of course, the processing time must be managed to avoid delay in the system response and to avoid bad user experience.

The system focuses on visual issues, but other senses, especially sound, can be explored together with the instructions. We identified in some AR approaches that the sound is an interesting auxiliary tool to get the user attention.

We did not apply visual attention models in the instruction system, which try to identify important areas on real scenes. The creation of automatic instructions can be significantly improved by the application of bottom-up and top-down visual attention models.

REFERENCES

ADDISON, M. A. and THIMBLEBY, H. W., 1995. Hypermedia manuals for interactive systems. Authoring and Application of Hypermedia-Based User-Interfaces, IEE Colloquium on, pp.5/1–5/4.

ADELSON, E.H. and ANANDAN, P. 1990. Ordinal characteristics of transparency. AAAI -Workshop on Qualitative Vision (Boston, MA, 1990), 77–81.

AGRAWALA, M., LI, W. and BERTHOUZOZ, F., 2011. Design principles for visual communication. Communications of the ACM, 54(4), p.60.

ANDERSEN, M., ANDERSEN, R., LARSEN, C., MOESLUND, T.B. and MADSEN, O., 2009. Interactive assembly guide using augmented reality. Advances in Visual Computing, pp.999–1008.

ANDERSON, F., GROSSMAN, T., MATEJKA, J. and FITZMAURICE, G., 2013. YouMove: Enhancing Movement Training with an Augmented Reality Mirror. Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST ’13, pp.311–320.

ASAI, K., KOBAYASHI, H. and KONDO, T., 2005. Augmented instructions - a fusion of augmented reality and printed learning materials. In Fifth IEEE International Conference on Advanced Learning Technologies (ICALT’05). IEEE Computer Society, pp. 213–215.

AUTODESK, 2015. Available in: http://www.autodesk.com.br/products/3ds-max/overview. Last view: January, 2015.

AVERY, B., SANDOR, C. and THOMAS, B.H. 2009. Improving Spatial Perception for Augmented Reality X-Ray Vision. 2009 IEEE Virtual Reality Conference. (Mar. 2009), 79–82.

AZUMA, R. and FURMANSKI, C., 2003. Evaluating label placement for augmented reality view management. In IEEE/ACM Internationl Symposium on Mixed and Augmented Reality. Tokyo, Japan: IEEE Comput. Soc, pp. 66–75.

BARTLETT, R., 2007. Introduction to Sports Biomechanics: Analysing Human Movement Patterns, 2 ed.: Routledge.

BERTALMIO, M., SAPIRO, G., CASELLES, V. and BALLESTER, C., 2000. Image inpainting. Proceedings of the 27th annual conference on Computer graphics and interactive techniques - SIGGRAPH ’00, pp.417–424.

BILLINGHURST, M., HAKKARAINEN, M. and WOODWARD, C., 2008. Augmented assembly using a mobile phone. In Mixed and Augmented Reality (ISMAR), 2008 7th IEEE International Symposium on. IEEE Computer Society, pp. 167–168.

BLUM, T., KLEEBERGER, V., BICHLMEIER, C. and NAVAB, N., 2012. mirracle: an augmented reality magic mirror system for anatomy education. 2012 IEEE Virtual Reality (VR), pp.115–116.

BMW SERVICE 2014. Available in: http://www.bmw.com/com/en/owners/service/augmented_reality_introduction_1.html. Acessed on: September 1th, 2014

BORJI, A. and ITTI, L., 2013. State-of-the-art in Visual Attention Modeling. IEEE Transaction on Pattern Analysis and Machine Intelligence, 35, pp.185–207.

BOUVIER-ZAPPA, S., OSTROMOUKHOV, V. and POULIN, P., 2007. Motion cues for illustration of skeletal motion capture data. Proceedings of the 5th international symposium on Non-photorealistic animation and rendering - NPAR ’07, p.133.

BRADSKI, G.R., 1998. Computer Vision Face Tracking For Use in a Perceptual User Interface. Intel TEchnology Journal Q2’98.

BRADSKI, G. and KAEHLER, A. 2008. Learning OpenCV, 1st ed. O'Reilly Media, Inc., 2008.

BRUCKNER, S. and GR, M.E., 2006. Exploded Views for Volume Data. IEEE Transactions on visualization and computer graphics, 12(5).

CANNY, J., A., 1986. Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679–698.

CHAVES, T., FIGUEIREDO, L., GAMA, AD., DE ARAUJO, C. and TEICHRIEB, V., 2012. Human Body Motion and Gestures Recognition Based on Checkpoints. Proceedings of Virtual and Augmented Reality (SVR), 2012 14th Symposium on, vol., no., pp.271,278, 28-31 May 2012 doi: 10.1109/SVR.2012.16

CHEN, C. 2005. Top 10 unsolved information visualization problems. IEEE computer graphics and applications. 25, 4 (2005), 12–6.

CHEN, J., CHEN, Y., GRANIER, X., WANG, J. and PENG, Q. 2011a. Importance-Driven Composition of Multiple Rendering Styles. 12th International Conference on Computer-Aided Design and Computer Graphics (CAD/Graphics) (2011), 79 –86.

CHEN, J., GRANIER, X. and LIN, N. 2009. On-Line Visualization of Underground Structures using Context Features. ACM Symposium on Virtual Reality Software and Technology (Hong Kong, 2009), 167–170.

CHEN, J., TURK, G. and MACINTYRE, B. 2011b. Painterly rendering with coherence for augmented reality. VR Innovation (ISVRI), 2011 IEEE International Symposium on 103–110.

COHEN, J., 2011. Statistical power analysis. Current Directions in Psychological Science, pp.98–101.

COHEN-OR, D., SORKINE, O., GAL, R., LEYVAND, T. and XU, Y.-Q. 2006. Color harmonization. ACM SIGGRAPH 2006 Papers on - SIGGRAPH'06 (New York, New York, USA, 2006), 624.

CRONBACH, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), pp.297–334.

DEITEL, H. M. and DEITEL, P. J. 2008. C++ How to Program. 6th ed. Deitel and Deitel, 2008.

DEITEL, H. M. and DEITEL, P. J. 2003. C# How to Program. Makron Books, 2003.

DUH, H.B.-L. and BILLINGHURST, M. 2008. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. (Sep. 2008), 193–202.

ELMQVIST, N. AND TSIGAS, P. 2008. A taxonomy of 3D occlusion management for visualization. IEEE transactions on visualization and computer graphics. 14, 5 (2008), 1095–109.

FARINAZZO M. V., DE PAIVA GUIMARAES, M., CORREA, A.G., MARTINS, V.F. and PAULO, S., 2013. Usability test for Augmented Reality applications. Computing Conference (CLEI), 2013 XXXIX Latin American, pp.1–10.

FEINER, S.K. and SELIGMANN, D.D., 1992. Cutaways and ghosting: satisfying visibility constraints in dynamic 3D illustrations. The Visual Computer, 8, pp.292–302.

FERWERDA, J. A. 2003. Three varieties of realism in computer graphics. Proceedings of SPIE, 290–297.

FIGUEIREDO, L., DOS ANJOS, R., LINDOSO, J., NETO, E., ROBERTO, R., SILVA, M. and TEICHRIEB, V. 2013. Bare hand natural interaction with augmented objects. 2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013, Adelaide, SA, 2013, pp. 1-6.

FISCHER, J., BARTZ, D. and STRABER, W. 2005. Stylized augmented reality for improved immersion. Virtual Reality, 2005. Proceedings. VR 2005. IEEE , vol., no., pp.195, 202, 12-16 March 2005.

FISCHER, J., HALLER, M. and THOMAS, B.H. 2008. Stylized depiction in mixed reality. The Int. Journal of Virtual Reality. 7, 4 (2008), 71-79.

FRINTROP, S., ROME, E. and CHRISTENSEN, H.I. 2010. Computational visual attention systems and their cognitive foundations: A survey. ACM Transactions on Applied Perception (TAP). 7, 1 (2010), 1–46.

FUKIAGE, T., OISHI, T. and IKEUCHI, K. 2012. Reduction of contradictory partial occlusion in mixed reality by using characteristics of transparency perception. IEEE International Symposium on Mixed and Augmented Reality (2012), 129–139.

FURMANSKI, C., AZUMA, R. and DAILY, M., 2002. Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and obscured information. In Proceedings of the International Symposium on Mixed and Augmented Reality. IEEE Computer Society, Washington, DC, USA, 215-.

GAMA, A. DA, CHAVES, T., FIGUEIREDO, L. and TEICHRIEB, V., 2012a. Guidance and Movement Correction Based on Therapeutics Movements for Motor Rehabilitation Support Systems. In 2012 14th Symposium on Virtual and Augmented Reality. IEEE Computer Society Washington, DC, USA, pp. 191–200.

GAMA, A. DA, CHAVES, T., FIGUEIREDO, L., TEICHRIEB, V., CARNEIRO, M., MARQUES-OLIVEIRA, D., BALTAR, A., CARDOSO, A. and MONTE-SILVA, K., 2012b. Ikapp – A Rehabilitation Support System using Kinect. In 14th Symposium on Virtual and Augmented Reality.

GIERLINGER, T., DANCH, D. AND STORK, A. 2010. Rendering techniques for mixed reality. Journal of Real-Time Image Processing. 5, 2 (Nov. 2010), 109–120.

GIMENO, J., MORILLO, P., ORDUÑA, J.M. and FERNÁNDEZ, M., 2013. A new AR authoring tool using depth maps for industrial procedures. Computers in Industry, 64(9), pp.1263–1271.

GOOCH, A.A., LONG, J., JI, L., ESTEY, A. and GOOCH, B.S. 2010. Viewing progress in non-photorealistic rendering through Heinlein's lens. Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering (2010), 165–171.

GOTO, M., UEMATSU, Y., SAITO, H., SENDA, S. and IKETANI, A., 2010. Task support system by displaying instructional video onto AR workspace. In 2010

IEEE International Symposium on Mixed and Augmented Reality. IEEE Computer Society, pp. 83–90.

GRUBER, L., KALKOFEN, D. and SCHMALSTIEG, D. 2010. Color harmonization for augmented reality. 2010 IEEE International Symposium on Mixed and Augmented Reality. (Oct. 2010), 227–228.

GUPTA, A., FOX, D., CURLESS, B. and COHEN, M., 2012. DuploTrack: a real-time system for authoring and guiding duplo block assembly. In Proceedings of the 25th annual ACM symposium on User interface software and technology (UIST'12). Cambridge, Massachusetts, USA: ACM New York, NY, USA, pp. 389–401.

HALLER, M. 2004. Photorealism or/and non-photorealism in augmented reality. Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry - VRCAI '04. (2004), 189.

HALLER, M. and SPERL, D. 2004. Real-time painterly rendering for MR applications. Proceedings of the 2nd international conference on computer graphics and interactive techniques in australasia and south east asia, (2004), 30–38.

HALLER, M., LANDERL, F. and BILLINGHURST, M. 2005. A loose and sketchy approach in a mediated reality environment. GRAPHITE '05 Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia. 1, 212 (2005), 371–379.

HANSEN, C., WIEFERICH, J., RITTER, F., RIEDER, C. and PEITGEN, H.-O. 2010. Illustrative visualization of 3D planning models for augmented reality in liver surgery. International journal of computer assisted radiology and surgery. 5, 2 (Mar. 2010), 133–41.

HENDERSON, S.J. and FEINER, S., 2009. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In 2009 8th IEEE International Symposium on Mixed and Augmented Reality. IEEE Computer Society Washington, DC, USA, pp. 135–144.

HENDERSON, S.J. and FEINER, S.K., 2011. Augmented reality in the psychomotor phase of a procedural task. In 2011 10th IEEE International Symposium on Mixed and Augmented Reality. IEEE Computer Society, pp. 191–200.

HERLING, J. and BROLL, W. 2012. PixMix: A Real-Time Approach to High-Quality Diminished Reality. IEEE International Symposium on Mixed and Augmented Reality 2012 (Atlanta, Georgia, 2012), 141–150.

HERLING, J. and BROLL, W. 2010. Advanced self-contained object removal for realizing real-time Diminished Reality in unconstrained environments. Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on (2010), 207–212.

HORIE, A., MEGA, S. and UEHARA, K., 2006. The interactive cooking support system in mixed reality environment. In IEEE International Conference on Multimedia and Expo. IEEE Computer Society, pp. 657–660.

HUGHES, J. F., DAM, A. V., MCGUIRE, M., SKLAR, D. F., FOLEY, J. D., FEINER, S. K. and AKELEY, K. 2013. Computer Graphics: Principles and Practice, 3rd edition, Addison-Wesley Professional.

ITTI, L., KOCH, C. and NIEBUR, E., 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Transaction on Pattern Analysis and Machine Intelligence, 20(11), pp.1254–1259.

IWAI, D., YABIKI, T. and SATO, K., 2013. View management of projected labels on nonplanar and textured surfaces. IEEE Transactions on Visualization and Computer Graphics, 19(8), pp.1415–1424.

IZADI, S., KIM, D., HILLIGES, O., MOLYNEAUX, D., NEWCOMBE, R., KOHLI, P., SHOTTON, J., HODGES, S., FREEMAN, D., DAVISON, A. and FITZGIBBON, A 2011. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In 24th annual ACM symposium on User interface software and technology (UIST '11). ACM New York, NY, USA, pp. 559–568.

JACOBS, K. and LOSCOS, C. 2006. Classification of illumination methods for mixed reality. Computer Graphics Forum. 25, 1 (Mar. 2006), 29–51.

JARUSIRISAWAD, S., HOSOKAWA, T. and SAITO, H. 2010. Diminished reality using plane-sweep algorithm with weakly-calibrated cameras. Progress in Informatics. 7 (Mar. 2010), 11-20.

KALKOFEN, D., MENDEZ, E. and SCHMALSTIEG, D. 2009. Comprehensible visualization for augmented reality. IEEE transactions on visualization and computer graphics. 15, 2 (2009), 193–204.

KALKOFEN, D., SANDOR, C., WHITE, S. and SCHMALSTIEG, D., 2011. Visualization Techniques for Augmented Reality. In B. Furht, ed. Handbook of Augmented Reality. New York, NY: Springer New York, pp. 65–98.

KALKOFEN, D., VEAS, E., ZOLLMANN, S., STEINBERGER, M. and SCHMALSTIEG, D., 2013. Adaptive ghosted views for augmented reality. In 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, pp. 1–9.

KHUONG, B.M., KIYOKAWA, K., MILLER, A., LA VIOLA, J.J., MASHITA, T. and

TAKEMURA, H., 2014. The effectiveness of an AR-based context-aware assembly support system in object assembly. 2014 IEEE Virtual Reality (VR), pp.57–62.

KIM, J. and SONG, S.Y., 1997. Instructional design guidelines for Virtual Reality in Classroom Applications. Trends and Issues in Educational Technology Research and Development.

KINECT. 2012. KINECT for Windows. Available at: http://www.microsoft.com/en-us/kinectforwindows/.

KNECHT, M., TRAXLER, C., PURGATHOFER, W. and WIMMER, M. 2011. Adaptive camera-based color mapping for mixed-reality applications. 2011 10th IEEE International Symposium on Mixed and Augmented Reality (Oct. 2011), 165–168.

KRUIJFF, E., SWAN, J. and FEINER, S. 2010. Perceptual issues in augmented reality revisited. Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on (2010), 3–12.

KYPRIANIDIS, J.E., COLLOMOSSE, J., WANG, T. and ISENBERG, T. 2013. State of the "Art": a taxonomy of artistic stylization techniques for images and video. IEEE transactions on visualization and computer graphics. 19, 5 (2013), 866-885.

LABORATORY ERGONOMICS, 2015. Anatomy and ergonomic fundamentals of human motion. Available in: http://www.chem.purdue.edu/chemsafety/safetyclass/injury/lecture/NIEHS.htm. Last view: January 2015.

LARMAN, C. 2004. Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and Iterative Development. 3th ed.

LEAO, C.W.M., LIMA, J.P., TEICHRIEB, V., ALBUQUERQUE, E.S. and KELNER, J. 2011. Altered reality: augmenting and diminishing reality in real-time. IEEE Virtual Reality (Singapore, 2011), 219–220.

LI, H. and NASHASHIBI, F. 2011. Multi-vehicle cooperative perception and augmented reality for driver assistance: A possibility to "see" through front vehicle. 14th International IEEE Conference on Intelligent Transportation Systems (Washington, DC, 2011).

LIU, C., PENG, Q. and WANG, X. 2011. Recent development in image completion techniques. Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on. 4, (2011), 756-760.

LIU, Y., QIN, X., XU, S., NAKAMAE, E. and PENG, Q. 2009. Light source estimation of outdoor scenes for mixed reality. The Visual Computer. 25, 5-7 (Mar. 2009), 637–646.

LIVINGSTON, M.A., AI, Z., KARSCH, K. and GIBSON, G.O. 2011. User interface design for military AR applications. Virtual Reality. 15, 2-3 (Jun. 2011), 175–184.

MADSEN, C.B., SORENSEN, M.K.D. and VITTRUP, M. 2003. The importance of shadows in augmented reality. 6th Annual International Workshop on Presence (Aalborg, Denmark, 2003), p. 21.

MAKITA, K., KANBARA, M. and YOKOYA, N., 2009. View management of annotations for wearable augmented reality. In IEEE International Conference on Multimedia and Expo - ICME, 2009. pp. 982–985.

MARCINČIN, J., BARNA, J., JANÁK, M., MARCINCINOVA, L.N. and FEČOVÁ, V., 2011. Utilization of open source tools in assembling process with application of elements of augmented reality. In Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry. ACM New York, NY, USA, pp. 427–430.

MCCRUM-GARDNER, E., 2010. Calculations Made Simple. , 17(1), pp.10–14.

MENDEZ, E. and SCHMALSTIEG, D. 2009. Importance masks for revealing occluded objects in augmented reality. VRST'09 Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology (2009), 247-248. ACM New York, NY, USA.

MENDEZ, E., FEINER, S. and SCHMALSTIEG, D. 2010. Focus and context in mixed reality by modulating first order salient features. Smart Graphics (2010), 232–243.

MENK, C. and KOCH, R. 2011. Interactive visualization technique for truthful color reproduction in spatial augmented reality applications. IEEE International Symposium on Mixed and Augmented Reality (Bazel, Switzerland, 2011), 157-164.

MILGRAM, P. and KISHINO, F. 1994. A taxonomy of mixed reality visual displays. IEICE Transactions on Information and Systems. 12 (1994), 1–15.

MILLER, A., WHITE, B., CHARBONNEAU, E., KANZLER, Z. and Jr, J.J.L., 2012. Interactive 3D model acquisition and tracking of building block structures. IEEE transactions on visualization and computer graphics, 18(4), pp.651–659.

MITRA, N.J., YANG, Y.-L., YAN, D.-M., LI, W. and AGRAWALA, M., 2010. Illustrating how mechanical assemblies work. ACM Transactions on Graphics, 29(4), p.1.

MIYAWAKI, K. and SANO, M., 2008. A virtual agent for a cooking navigation. In 8th International Conference, IVA. Tokyo, Japan: Springer Berlin Heidelberg, pp. 97–103.

MOTOKAWA, Y. and SAITO, H., 2007. Support system for guitar playing using augmented reality display. In Proceedings - ISMAR 2006: Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality. pp. 243–244.

NAWAHDAH, M. and INOUE, T., 2011. Helping physical task learning by automatic adjustment of a virtual teacher's rotation angle. Proceedings of the 2011 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD), pp.710–715.

NEUMANN, U. and MAJOROS, A., 1998. Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180), pp.4–11.

NIENHAUS, M. and DOLLNER, J., 2005. Depicting dynamics using principles of visual art and narrations. Computer Graphics and Applications, 25(June), pp.40–51.

NILSSON, S. and JOHANSSON, B., 2008. Acceptance of augmented reality instructions in a real work setting. Proceedings of ACM CHI 2008 Conference on Human Factors in Computing Systems, 2, pp.2025–2032.

NILSSON, S. and JOHANSSON, B., 2007. Fun and usable: augmented reality instructions in a hospital setting. In OZCHI '07 Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces. ACM New York, NY, USA, pp. 123–130.

NOH, Z. and SUNAR, M.S. 2009. A review of shadow techniques in augmented reality. 2009 Second International Conference on Machine Vision. (2009), 320–324.

NOWROUZEZAHRAI, D., GEIGER, S., MITCHELL, K., SUMNER, R., JAROSZ, W. and GROSS, M. 2011. Light factorization for mixed-frequency shadows in augmented reality. 2011 10th IEEE International Symposium on Mixed and Augmented Reality. (Oct. 2011), 173–179.

PADILHA A. and TEICHRIEB, V. 2014. Motion detection based ghosted views for occlusion handling in augmented reality. 17th IEEE International Symposium on Mixed and Augmented Reality.

PADILHA, A., ROLIM, C. and TEICHRIEB, V., 2013. The ghosting technique applied to augmented reality visualization. In XV Symposium on Virtual and

Augmented Reality (SVR). pp. 159–166.

PUIG, J., PERKIS, A., LINDSETH, F. and EBRAHIMI, T., 2012. Towards an efficient methodology for evaluation of quality of experience in Augmented Reality. Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on, pp.188–193.

PATHOMAREE, N. and CHAROENSEANG, S., 2005. Augmented reality for skill transfer in assembly task. In ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005. IEEE Computer Society, pp. 500–504.

PETERSEN, N. and STRICKER, D., 2012. Learning task structure from video examples for workflow tracking and authoring. In IEEE International Symposium on Mixed and Augmented Reality. Atlanta, Georgia: IEEE Computer Society, pp. 237–246.

PETERSEN, N., PAGANI, A. and STRICKER, D., 2013. Real-time modeling and tracking manual workflows from first-person vision. 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp.117–124.

RAGHAVAN, V., MOLINEROS, J. and SHARMA, R., 1999. Interactive evaluation of assembly sequences using augmented reality. IEEE Transaction on Robotics and Automation, 15(3), pp.435–449.

REINERS, D., STRICKER, D., KLINKER, G. and MÜLLER, S., 1998. Augmented reality for construction tasks: Doorlock assembly. In Proceedings of the international workshop on augmented reality: placing artificial objects in real scenes: placing artificial objects in real scenes. A. K. Peters, Ltd., pp. 31–46.

REINHARD, E., AKYUZ, A. and COLBERT, M. 2004. Real-time color blending of rendered and captured video. Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) (2004), 1–9.

ROBERTO, R.A., Freitas, D.Q. De, Simões, F.P.M. and Teichrieb, V., 2013. A dynamic blocks platform based on projective augmented reality and tangible interfaces for educational activities. Proceedings - 2013 15th Symposium on Virtual and Augmented Reality, SVR 2013, 4(2), pp.1–9.

ROGERS, Y., SHARP, H. and PREECE, J., 2005. Interaction Design: beyond human-computer interaction. 1st edition, John Wiley & Sons, Inc.

ROLIM, C. and TEICHRIEB, V. 2012. A viewpoint about diminished reality: is it possible remove objects in real-time from scenes? 2012 14th Symposium on Virtual and Augmented Reality. (May. 2012), 141–146.

ROLIM, C., SCHMALSTIEG, D., KALKOFEN, D. and TEICHRIEB, V., 2015.

[POSTER] Design Guidelines for Generating Augmented Reality. In IEEE International Symposium on Mixed and Augmented Reality. Fukuoaka, Japan.

SANDOR, C., CUNNINGHAM, A., DEY, A. and MATTILA, V.-V. 2010. An augmented reality x-Ray system based on visual saliency. 2010 IEEE International Symposium on Mixed and Augmented Reality. (Oct. 2010), 27–36.

SANTOS, A.L. DOS, LEMOS, D., LINDOSO, J.E.F. AND TEICHRIEB, V. 2012. Real-time ray tracing for augmented reality. 2012 14th Symposium on Virtual and Augmented Reality. (May. 2012), 131–140.

SAWANT, N. and MITRA, N.J. 2008. Color harmonization for videos. 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing (Dec. 2008), 576–582.

SAYIM, B. and CAVANAGH, P. 2011. The art of transparency. I-Perception. 2, 7 (Jan. 2011), 679–96.

SELIGMANN, D.D. and FEINER, S., 1991. Automated generation of intent-based 3D illustrations design rules: mapping intent to stylistic. In SIGGRAPH '91 Proceedings of the 18th annual conference on Computer graphics and interactive techniques. New York, NY, USA: ACM Press, pp. 123–132.

SEO, B.-K., LEE, M.-H., PARK, H. and PARK, J.-I. 2008. Projection-based diminished reality system. 2008 International Symposium on Ubiquitous Virtual Reality. (Jul. 2008), 25–28.

SHANMUGAM, P. 2015. UX & Virtual Reality - Designing for interfaces without Screens. Available in: http://www.uxness.in/2015/08/ux-virtual-reality.html. Last view: March, 2015.

SHREINER, D., SELLERS, G., KESSENICH, J. and LICEA-KANE, B., 2013. OpenGL programming guide: The Official Guide to Learning OpenGL, Version 4.3, 8/E . Addison-Wesley Professional.

SIGG, S., FUCHS, R., CARNECKY, R. and PEIKERT, R., 2012. Intelligent cutaway illustrations. 2012 IEEE Pacific Visualization Symposium, pp.185–192.

SILVA, S., SOUSA SANTOS, B. and MADEIRA, J. 2011. Using color in visualization: A survey. Computers & Graphics. 35, 2 (Apr. 2011), 320–333.

SUGANO, N., KATO, H. and TACHIBANA, K. 2003. The effects of shadow representation of virtual objects in augmented reality. Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on (2003), 76–83.

SZELISKI, R. 2010. Computer vision: algorithms and applications. Springer, New York.

TANG, A., OWEN, C., BIOCCA, F. and MOU, W., 2003. Comparative effectiveness of augmented reality in object assembly. In Proceedings of the conference on Human factors in computing systems - CHI '03. New York, New York, USA: ACM Press, pp. 73–80.

TANG, R., TANG, A., SCOTT, X.Y., BATEMAN, S. and JORGE, J., 2015. Physio@Home: Exploring visual guidance and feedback techniques for physiotherapy exercises. In 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). pp. 4123–4132.

TATZGERN, M., KALKOFEN, D. and SCHMALSTIEG, D., 2013. Dynamic compact visualizations for augmented reality. 2013 IEEE Virtual Reality (VR), pp.3–6.

THE NEW YORKER TIMES MAGAZINE, 2012. Available in: http://www.nytimes.com/2012/01/08/magazine/how-yoga-can-wreck-your-body.html?_r=0. Last view: September, 2015.

TORY, M. and MOLLER, T. 2004. Human factors in visualization research. Visualization and Computer Graphics, IEEE Transactions on. 10, 1 (2004), 72–84.

TSUDA, T., YAMAMOTO, H., KAMEDA, Y. and OHTA, Y., 2005. Visualization methods for outdoor see-through vision. Proceedings of the 2005 international conference on Augmented tele-existence - ICAT '05, p.62.

URATANI, K., MACHIDA, T., KIYOKAWA, K. and TAKEMURA, H., 2005. A study of depth visualization techniques for virtual annotations in augmented reality. In IEEE Virtual Reality Conference. IEEE Computer Society, pp. 295–296.

VIOLA, I., KANITSAR, A. and GROLLER, M.E., 2005. Importance-driven feature enhancement in volume visualization. IEEE transactions on visualization and computer graphics, 11(4), pp.408–418.

WANG, S., CAI, K., LU, J., LIU, X. and WU, E. 2010. Real-time coherent stylization for augmented reality. The Visual Computer. 26, 6 (2010), 445–455.

WANG, Y.-S., CHEN, C.-M., HONG, C.-M. and TSAI, Y.-N., 2013. Interactive augmented reality game for enhancing library instruction in elementary schools. 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops, pp.391–396.

WARD, J, G., 1989. The RADIANCE lighting simulation and rendering system. 21st annual conference on Computer graphics and interactive techniques, pp.459–472.

WEBEL, S., BOCKHOLT, U., ENGELKE, T., GAVISH, N., OLBRICH, M. and PREUSCHE, C., 2013. An augmented reality training platform for assembly and maintenance skills. Robotics and Autonomous Systems, 61(4), pp.398–403.

WITHER, J., DiVERDI, S. and HÖLLERER, T., 2009. Annotation in outdoor augmented reality. Computers & Graphics, 33(6), pp.679–689.

WONG, B., SHIH, K., LIANG, C. and CHEN, H.H. 2012. Single image realism and recoloring by color compatibility. IEEE Transaction on Multimidia. 14, 3 (2012), 760–769.

XNA GAME STUDIO, 2015. Available in: http://msdn.microsoft.com/en-us/library/bb200104.aspx. Last view: January, 2015.

YAMADA, T., KOJIMA, H. and YAJIMA, H., 1997. Human interface design methodology for electronic manual system and its application. 1997 IEEE 6th International Conference on Emerging Technologies and Factory Automation Proceedings, EFTA '97.

YOGA, I. 2013. Available in: http://chereeyoga.net/tag/sequences/. Last view: January, 2015.

YEOH, R.C. and ZHOU, S.Z. 2009. Consistent real-time lighting for virtual objects in augmented reality. 2009 8th IEEE International Symposium on Mixed and Augmented Reality. (Oct. 2009), 223–224.

ZAUNER, J., HALLER, M. and BRANDL, A., 2003. Authoring of a mixed reality assembly instructor for hierarchical structures. In Proceedings of the Second IEEE and ACM International Synposium on Mixed and Augmented Reality. IEEE Computer Society, pp. 237–246.

ZHU, J., PAN, Z., SUN, C., and CHEN, W. 2010. Handling occlusions in video-based augmented reality using depth information. Computer Animation Virtual Worlds 21, 5 (September 2010), 509-521.

ZOKAI, S., ESTEVE, J., GENC, Y. and NAVAB, N. 2006. Multiview paraperspective projection model for diminished reality. The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings. (2006), 217–226.

ZOLLMANN, S., KALKOFEN, D., MENDEZ, E. and REITMAYR, G. 2010. Image-based ghostings for single layer occlusions in augmented reality. Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on (Seoul, Korea, 2010), 19–26.