Overview

Title: Personal Perception and Interaction with Mixed Reality

1. *Introduction*: The history of MR, focusing on how to improve the perception of the digital information and the corresponding interactive methods. objectives and requirements in different stages and how the technology develops.

2. *The state of art* The current states: some famous MR system, the main contributions of the MRs and how to improve Personal Perception and Interaction.

3. *Personal Perception with Magic Mirror*: How to improve the perception in the Magic Mirror framework.

    a) *Magic Mirror framework* conception, software and hardware

    b) *Accuracy of Registration* improve the perception of the MR view via accurate registration. Anatomy learning, personal information (gender, age, body shape)

    c) *Interactive MR* Make the user believe the virtual element is a part of his or her own body. (Organ game and the Muscle learning)

4. *Personal Pointing Interaction*: The interaction with MR via pointing gesture

    a) *Calibration and Recovery of Pointing gesture* The PAST method and the evaluation.

    b) *Interaction with Multi-screen*

    c) *Augmented Human to access the real world* The case without depth info. Natural pointing interaction for optical see-through AR to help wearable computer to fetch the information

5. *Collaborative Mixed Reality* combine the technology for perception and interaction (Maybe Collaborative User interface or platform.)

    a) *Magic mirror with pointing gesture interaction* One patient <-> one Kinect and one display doing the rehabilitation exercise. One nurse with pointing gesture device: to control all the exercises. Parents monitor children's magic mirror teacher monitor student's magic mirror

    b) *Shared the view* one user perform pointing, another user really see the pointing gesture in his or her own view. to collaborate with each other. For teaching and education.

1 Introduction

[Generic blabla about Augmented reality, perception and interaction.] Information: physical world and real objects => digital information => text info => graphic display 2D => 3D, go outside from computer and shown in the real world. natural perception.

Learning => personal knowledge network construction and personal perception is become important.

VR => AR => MR and

The history of MR, focusing on how to improve the perception of the digital information and the corresponding interactive methods. objectives and requirements in different stages and how the technology develops.

interaction with real object by touch and control

WIMP and cursor and mouse and keyboard.

computer vision algorithm and natural interaction based on the depth image and human tracking

smart agent based on machine learning to help computer understand the user

interaction with digital and physical element more and more like superman

AR and MR for education.(one section) see the ISMAR2013-5pages document section 2

1.1 Motivation

Talk about the general motivation of this work.

new conception and method to improve the perception with personal information MR or AR registration with help of personal information and personal interaction with the MR view

1.2 Objectives

1.3 Contributions

2 Personal Perception with Magic Mirror

The reason and advantage of the magic mirror conception. The history of magic mirror conception

2.1 Magic Mirror Conception

The General Medical Council recently proposed standards for effective teaching and learning of medical students [CG09]. They stated that: "...medical schools should take advantage of new technologies. . . . to deliver teaching." Augmented reality research has matured to a level that its applications can now be found in both mobile and non-mobile devices [Bac+14] and research on AR has also demonstrated its extreme usefulness for increasing student motivation in the learning process [Cha+14; Di +13]. Providing adequate learning experience to different learners is a challenging issue as the learning system generally does not adapt content to suit individual learner needs. Personalization for promoting a multimodal learning environment is also a growing area of interest, such as the development of user modeling and personalized processes which place the student at the center of the learning development.

2.1.1 System

Our system prototype has a mirror-like effect to the user by projecting a 'looking glass' on the body. The prototype of the magic mirror is largely based on a software framework that has been developed for HMD-based AR. The system can currently display the skeleton of the user, rendered from pre-operative CT data. The mirror tracks users' movements using a depth camera and an algorithm to detect the pose of the user from the depth image. This is realized using the Microsoft Kinect which was originally developed to allow controlling computer games by motion. By using the magic mirror metaphor, the user is led to believe that he or she is able to look inside their own body. At the same time, radiological information (CT, MRI data and a fully segmented dataset of cross-sectional photographs of the human body) are displayed in real-time. The user is able to select slices from these dataset by hand movement [Blu+12; Nav+12].

The current system allows visualization of static anatomy on the user and offers a simple user interface to select CT, MR or photographic slices. Demos of the prototype have been given on various public occasions. This exposure to the public resulted in a several contacts with medical doctors (MD) from physiotherapy, anatomy, radiology, sports medicine, children's medicine, and with medical teacher and science centers who are interested in using such a system in education of medical students and pupils. While discussions with these contacts were encouraging, most MDs asked for precision.

Hardware setup

Software framework

The interactive anatomy learning system which is presented in this paper is an augmented reality magic mirror and has primarily been developed for medical anatomy education, museums, and exhibitions. It focuses on bones and important organs of the thorax and abdomen. Fig.**??** depicts the general view of the system, which includes a large display device and a Microsoft Kinect. The Microsoft Kinect contains both color and depth camera and is sold as an add-on for the Xbox 360 video game console, taking gestures, color image, body movement, and voice as game input. The user's skeleton and personal information can be generated from the Kinect sensor. The prototype of our magic mirror is largely based on a software framework that has been developed for HMD-based AR. Our system prototype has a mirror-like effect to the user by projecting a 'looking glass' on the user body. The system can currently provide AR in-situ visualization, render skeleton from an original CT-volume, and showcase 3D models of organs.

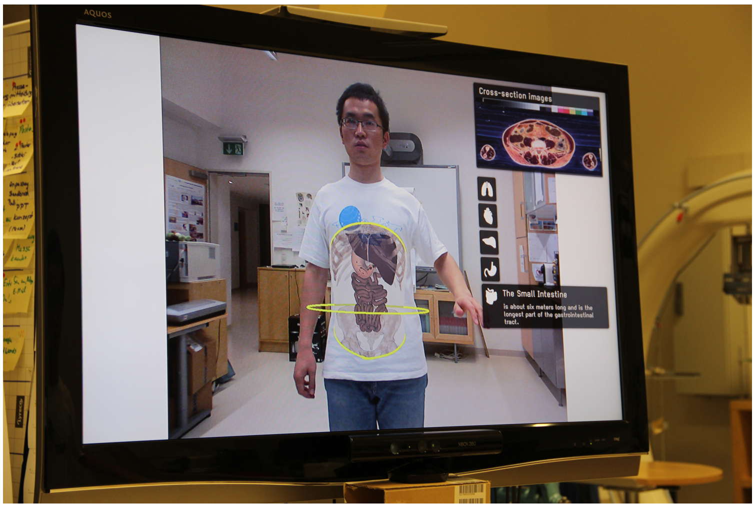


Fig. 2.1 Personalized magic mirror system for anatomy learning. A sensor tracks the user positions in real-time, and contextual in-situ visualization algorithms enable the augmentation of organs and muscles directly onto the user body.

2.1.2 Natural perception and interaction

AR in-situ visualization of human anatomy

For a personalized visualization of organs we employ the concept of a magic mirror. The camera image is flipped horizontally and shown on the screen such that the user has the impression of standing in front of a mirror. The system tracks all user movements using a depth camera and an algorithm to detect the pose of the user from the depth image. This is realized using the Microsoft Kinect which was originally developed to allow controlling computer games by motion. Then, virtual objects can be added to the image of the real scene. By using the magic mirror metaphor, the user is led to believe that he or she is able to look inside their own body. At the same time, radiological information (CT data and a fully segmented dataset of cross-sectional photographs of the human body) are displayed in real-time [Blu+12].

Prior to correctly using our AR magic mirror system, the users stand in front of the monitor and we displayed virtual marks near the five bone landmarks. The users are asked to interactively adjust the positions of the five marks to fit their own bone positions. In addition, the exact locations in the VKH CT dataset of the five selected bone landmarks are known. A linear interpolation was executed to estimate the torso point (i.e. a 6th landmark) in the CT volume to improve the overlay. Then the scale factors and transformation matrix were computed to render the anatomical image onto the user's body. These landmark positions allow the deformation and interpolation of the medical data correctly within the magic mirror and onto the human body, resulting in a more precise augmentation. A user study involving surgeons and anatomy experts confirmed our findings and will be presented in Section 3.

Natural User interaction

2.1.3 Potential application with magic mirror

Medical education AR rehabilitation Patient positioning

2.2 Personal Registration of Magic Mirror

2.2.1 Personal Information

We augment a personalized visualization of a CT dataset onto the user. However as full CT scans of any user are generally not available, we use the Visible Korean Human dataset (VKH), which consists of a CT scan, an MR volume and a photographic volume which has been acquired by stacking up cry sections. To allow a correct augmentation of the CT data, the gender, age, body size and pose of the user has to be detected. This is performed based on the color image using the OpenBR library [Klo+13] and the depth image using the NITE skeleton tracking software. The corresponding CT volume is chosen and scaled to the size of the user and augmented onto the user body. For visualization of the bones a transfer function is used as bones can be distinguished easily in the CT volume based on their voxel intensities. For the organs a segmentation of the VKH is used. The augmentation uses contextual in-situ visualization such that the virtual objects are only shown through a circular window. This leads to a better perception of depth, compared to a simple augmentation of the whole CT. The user could naturally turn around their body to check CT volume from different viewpoints and put their right hand at different heights to select a body plane with corresponding anatomy.

2.2.2 Personal Registration

We presents a general method to interactively improve and correct the Kinect skeleton for anatomy education purposes. We believe that our general method can be applied to projectors or other sensors as well for augmented reality. A thorough validation of our method demonstrated improved precision of anatomical landmarks and opens the avenue to future improvements in medical education. Together with the ISMAR community, we hope to initiate such discussions in integrating exciting user-interaction and gaming concepts within our system.

Our AR magic mirror relies on Kinect sensor which offers an imprecise skeleton tracking output (see Fig.??). This limits the precision of our magic mirror augmentations offering users false anatomical positions overlaid onto their body, resulting in a poor medical learning environment. Alternatively, had we considered projectors to display human anatomy directly on a user's body the same inaccuracies would exist[Sun13].

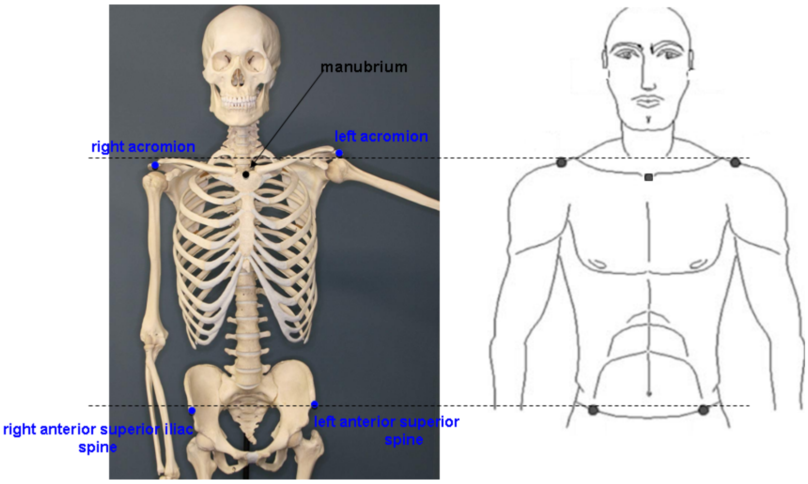


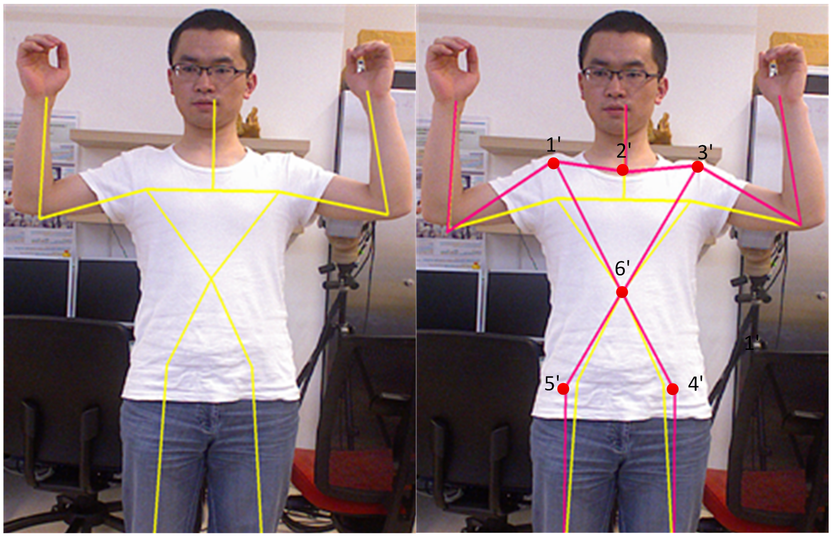
Fig. 2.2 Selected anatomical points for Kinect skeleton improvement and subsequent CT warping and interpolation.

The goal of this method is to propose a more precise user-specific learning environment. Together with orthopedic surgeons we have defined anatomical bone landmarks: (i) which are correctly identified in medical data such as CT and (ii) which users can touch easily on their body while standing in front of any sensor. These landmark positions allow the deformation and interpolation of the medical data correctly within the magic mirror and onto the human body, resulting in a more precise Kinect skeleton and augmentation. A user study involving surgeons and anatomy experts confirm our findings.

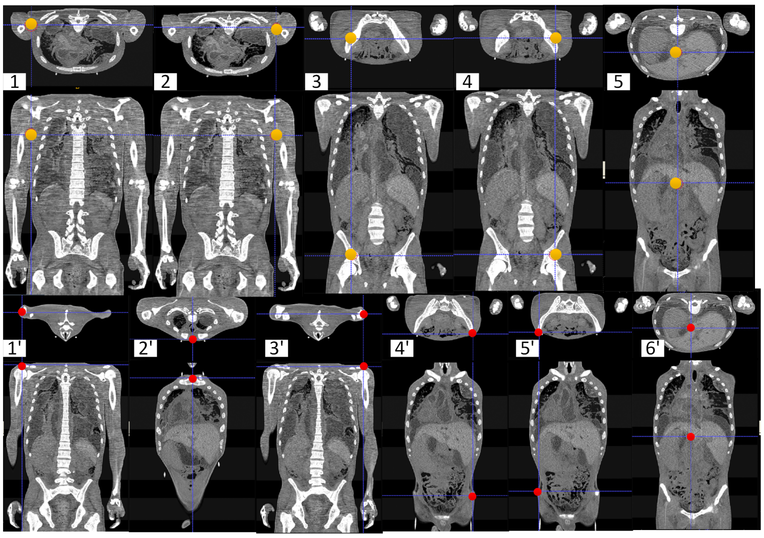
The skeleton output from Kinect limits the precision of the AR magic mirror applications. Thus, our magic mirror augmentations would contain errors and users would easily distinguish anatomical offsets on their body resulting in a poor anatomy learning environment. Alternatively, we had considered projectors to display human anatomy and along with their existing limitations the same inaccuracies would exist. Our aim was then to propose a method to correct for such overlay inaccuracies that could be translated to any magic mirror system worldwide. Together with orthopedic surgeons, we have defined five bone landmarks that (i) users can easily identify and touch while standing in front of the Kinect, and (ii) that are accurately and easily identified in medical data such as CT [Ma +13]. These are: left and right acromion, left and right anterior superior iliac spine, and the manubrium (see Fig.??).

Anatomical bone landmarkers

Together with orthopedic surgeons, we have defined five bone landmarks that can easily be touched on the human body. These are: left and right acromion, left and right anterior superior iliac spine, and the manubrium. Subjects are positioned in front of the Kinect sensor and asked to interactively adjust the positions of five landmarks



(a) Improvement of Kinect skeleton



(b) Bone landmarker in CT volume

Fig. 2.3 (a-b) The inaccurate Kinect skeleton, in yellow, compared to the improved Kinect skeleton positions in red. (c) CT data from the Visible Human Korean showing the 5 landmarks from Kinect skeleton. (d) CT data showing correctly the 5 landmarks + interpolated torso landmark using our method.

Tab. 2.1 Precision (in cm) of magic mirror system based on anatomical offsets

Anatomy	Offset(Mean± STDev)
anterior superior iliac spine	0.67 ± 0.52
manubrium	0.67 ± 0.75
heart	1.17 ± 1.60
liver	1.33 ± 1.21

Improment of kinect skeleton

Figure 3 shows a comparison between the traditional Kinect skeleton and its proposed improvement. The first row depicts visually the exactness of the new skeleton. The second row depicts the skeleton landmarks directly on CT data. We observe that the shoulder and anterior superior spine are inaccurate in the images. The last row depicts the improved landmark positions within CT as well. Transverse and sagittal CT slices of the visible human Korean are seen respectively in rows 2 and 3. In Figure 3a, the following scale factors were computed for the magic mirror augmentation:

Evaluation

As an augmented reality anatomy learning application, both accuracy and system usability are very important prior to its translation in classroom. Firstly, we undertook one user study with particular users having expert anatomy knowledge to evaluate if this system is precise enough for anatomy learning. Secondly, another user study involving first year medical students took place to verify the learning potential and acceptability of our technology as a compliment to atlas textbooks in classroom.

Assessing the magic mirror system precision and usability Participants: Seven participants were included in this study (two surgeons and five final year undergraduate medical students). Analysis: a Likert scale was used which is a type of psychometric response scale often used in surveys and the most widely used scale in survey research. When responding to a Likert questionnaire item, respondents specify their level of agreement to a statement. The format of our 5-pt Likert was: (1) strongly disagree, (2) disagree, (3) neither agree nor disagree, (4) agree, (5) strongly agree. To assess the precision of our personalized magic mirror we asked the participants to interact with the system platform which integrates user-specific anatomical landmark selection. Participants were asked to provide an estimated numerical offset, if any, on how far specific bone landmarks or organs were with respect to their own body. For this, they interacted with the magic mirror window, CT data, and used their own medical knowledge and expertise for judgment. CT data was displayed in an interface depicting both transverse and sagittal planes, and participants would quantify the offsets. If needed, a ruler was provided to assist them. The anatomical targets during evaluation were defined as: the anterior superior iliac spine, manubrium, heart, and liver. Results from this exercise are shown in Tab.2.1, with offsets measured in centimeters. Results from the user study show that the precision of user-specific learning environment is on average 0.96cm.

The seven participants were then asked to judge the usability of the AR magic mirror system by responding to the following questions: (i) is the overlay accurate w.r.t human body (ii) is the user interface easy to use, (iii) is it fun to play, (iv) can it be used for medical education, and

Tab. 2.2 Likert scale results regarding magic mirror usability

	Mean$\pm$ STDev
is the augmented reality overlay accurate w.r.t human body	4.00 ± 0.89
is the user interface easy to use	3.67 ± 1.03
is it fun to play	4.50 ± 0.55
can it be used for medical education	4.17 ± 0.75

(v) would it have stronger impact for medical education learning? The Likert scale results for the first four questions are shown in Tab.tb:3-PRMM:results2. For the last question regarding the impact of our technology, there was a unanimous response that the AR magic mirror system should be considered as a potential platform to complement existing anatomy learning tools inside anatomy classrooms.

Discussion

The precision of our method is visually demonstrated in Figure 3c-d and Fig.**??**. We observe that the acromion and anterior superior iliac spine, using the traditional Kinect skeleton, is not positioned correctly within the CT data compared to the modified Kinect skeleton version (1-2 vs. 1'-2'; 3-4 vs. 3'-4'). The orthopedic surgeons participating in our study confirmed this. Results from a user study show the impact of interactively improving the Kinect skeleton to increase precision for a better visualization of anatomy. The offsets of specific anatomical landmarks decreased significantly. The following comments were collected:

1. the precision improved; making the user touch anatomical landmarks is cool since this is the way it is done in clinic…
2. two additional landmarks easily accessible are the sternum and bottom of rib cage…
3. make the magic mirror circle bigger for larger anatomy…
4. voice command is a good idea but it is sensitive to the surroundings voice
5. interactions between observers …
6. could introduce the female CT visible human Korean…
7. could use a healthy patient CT or other modality…
8. could make the CT slices bigger on the screen…

One appealing feature of the system is that with the Kinect we are using inexpensive standard hardware. In the future such a system could be made available to students or patients who have to do rehabilitation exercises at home. In addition to the full system using a large screen and a screen stand we also want to evaluate the benefit of making the system accessible to students when they are at home and at any time. We want to do this for two different reasons. First, there is a trend toward competency-based education in medicine. Instead of defining a curriculum, learning outcomes are defined. Students have to fulfill these learning

Fig. 2.4 Figure 4: The magic mirror before (top) and after (bottom) the Kinect skeleton adjustment. Column 2 depicts the bottom of the rib cage being positioned correctly after skeleton correction.

outcomes. The advantage of competency-based education is that all students will have the same competency in the end. A student who is less skilled has to take more time to learn than a student who already has good skills. One requirement for this is that the students have to be able to educate themselves until they reach the required competency. We plan to use the AR magic mirror system both to allow them to do training and to test whether a learning outcome has been met. The magic mirror has to be made easily available to them so that they can use it for training at any time. The second reason to develop a distributed system which can easily be used by students is to collect user statistics.

Many medical education systems are web-based and allow accessing them from every computer. However current systems do not use high-end visualization and input devices like the Kinect. In the future when technologies like HTML5 and WebGL have matured it can be imagined that a system like the magic mirror could be implemented using web technology such that it can run on any computer. However, at the moment the use of web technology would be a significant limitation. We are using very large datasets, high-end visualization, GPU computing and gesture-based interaction.

It was originally suggested to us that our system would have much more of an impact for medical education learning if it were to be translated today in medical schools and anatomy classrooms. As such, with the help of our co-author and anatomy professor, we deliver the improved AR magic mirror to two anatomy classes within the Anatomy Department of the Ludwig Maximilian University (LMU) Medical School, Munich, Germany. The anatomy professors made it clear that the AR magic mirror system is exciting, but has to go beyond state-of-the-art technology to be truly useful for education. Our participants stressed the importance of visualization of anatomy that can change dynamically resulting from the actions of the user moving the body, and also for visualization of muscles. Furthermore, more advanced user interactions like the use of gaming elements would be required to make the use of the system for learning tasks more interesting.

2.3 Interactive Mixed Reality

2.3.1 Personalized natural interactin & gaming

Gesture-based interaction : Medical volumes are usually visualized by showing slices that are aligned with the axes of the volume. A volume can be seen as a stack of transverse slices starting from the top and going to the bottom. When the system is in the transverse slice mode, the user could move their hand up or down to choose the interesting slice image, respectively left or right for the sagittal slice mode. The current slice is depicted on the right part of the monitor while a yellow circle is augmented onto the user depicting the slice plane. The system can easily switch slices between the CT and the photographic volume.

Organ explosion effect : In the AR learning environment, organ selection is a challenging task as some organs hide behind others. The user can't directly position their hand into their bodies and to select the organ of choice. In engineering design, mechanical parts are drawn with the innermost part at the center, while the others are moved some fixed distance outwards to display all parts of the assembly that would otherwise be hidden. Inspired by this principle we designed and developed an organ explosion effect (see Figure 3-left), which separates the organs. The organ explosion selection employs a two-hand interaction method. Using the left hand, the user is able to focus the height for the section they are interested in. Organs at approximately the same height are then projected outwards. Using a spherical projection, the organs are moved outward, creating the illusion of seeing them 'fly out' in front of the body. Lines are drawn to indicate the original position of the organs. After separating the organs from each other, it is easy for the user to select the organ of interest using their right hand. Another functionality of the organ explosion effect is allowing the user to rotate and observe the 3D organ models and perceive the spatial relationship between them.

Serious gaming : Health education is a particular area where serious gaming is suited [Aub+12; Han+12]. We developed a game for general users to learn basic anatomy called Whack an Organ (see Figure 3-right). Our game is based on the idea of the classic arcade games at carnivals, 'Whack-a-Mole'. We implement a similar game idea using our system framework. It is a combination of Whack-a-Mole and classical quiz games: the user is presented with questions regarding human anatomy, and the answers are always organs or organ systems. To answer the question, the user has to point to the location of the organ directly on their body. The game then decides if the location is correct and awards points if so. Questions can come from different question sets with varying difficulty, ranging from simple location questions: where is the liver?, to more complex knowledge questions: which organ is infected if you have Hepatitis?