Vision-Based User Interfaces for Health Applications: A Survey

Alexandra Branzan Albu

Dept. of Electrical and Computer Engineering, University of Victoria (BC), Canada
aalbu@ece.uvic.ca

Abstract. This paper proposes a survey of vision-based human computer interfaces for several key-fields in health care: data visualization for image-guided diagnosis, image-guided therapy planning and surgery, the operating room, assistance to motor-impaired patients, and monitoring and support of elderly. The emphasis is on the contribution of the underlying computer vision techniques to the usability and usefullness of interfaces for each specific domain. *It is also shown that end-user requirements have a significant impact on the algorithmic design of the computer vision techniques embedded in the interfaces.*

1 Introduction

The field of computer vision focuses on the development and implementation of algorithms which allow computers to "understand" image and video data at various levels depending on the task at hand. Task-oriented image "understanding" may offer assistance to human perception, cognition and decision-making, such as in computer-aided diagnosis systems, or may enable more natural ways for human-computer interaction (HCI) in perceptual interfaces and in pervasive computing systems. Such vision-based technologies find promising applications in several areas of health care, including but not limited to image-based diagnosis and therapy planning, minimally invasive surgery, assistance and support for people with disabilities and elderly.

However, HCI design for medical applications is a difficult problem. Gosbee and Ritchie [1] built a hierarchical model of clinician acceptance of technology in an attempt to identify why physicians and other care providers are reluctant to introduce new HCI technology in their daily work routine. Considering this model, it is expected that a successful integration of computer vision algorithms in health-related HCI should consider both user-centered and task-based design paradigms. In return, these paradigms influence the basic assumptions as well as the algorithmic development of computer vision techniques.

This paper presents a survey of vision-based HCIs for several key-fields in health care: data visualization for image-guided diagnosis, therapy planning and surgery (section 2), the operating room (section 3), assistance to motor-impaired patients (section 4), and monitoring and support of elderly (section 5). The emphasis is on the contribution of the embedded computer vision techniques to the usability and usefullness of interfaces for each specific domain. Section 6 presents a summary of our survey and draws conclusions.

2   Vision-Based Interfaces for Enhanced Data Visualization

Computer vision algorithms for medical image analysis form four main groups with different scopes, as follows: a) filtering for image enhancement; b) segmentation for object delineation; c) analysis for feature extraction; d) image registration for multi-modal data fusion. Recently, applications for advanced medical data visualization such as virtual or augmented reality systems integrate computer vision with computer graphics. While computer vision deals with extracting relevant information from images (e.g. object boundaries, interstructural distances), computer graphics focuses on image synthesis for creating realistic and manipulable 3D objects representing anatomical structures of interest. Computer vision and computer graphics techniques are strongly interconnected in user interfaces (UI) for data visualization.

HCI for computer-aided diagnosis and therapy planning in clinical environments. In this context, a particular attention is to be directed towards the algorithmic design of segmentation techniques in order to meet end-user requirements collected from radiologists. Yet, relatively recent surveys on interactive and fully automated methods for medical image segmentation [2, 3] do not discuss the relevance of end-user requirements for segmentation design. One may reach the surprising conclusion that the development of most segmentation methods is not driven by end-user requirements; most developers focus on improving accuracy, precision and computational speed, as well as on reducing the amount of manual interaction. Widely recognized validation protocols for medical image segmentation [4, 5] perform a strict scientific comparison of the performance of automatic/semi-automatic methods against manual expert segmentations. Few methods search clinical feedback and conduct a usability study. Such a method is proposed by O'Donnell et al [6], and consists in a user-steered segmentation algorithm based on the interactive livewire paradigm [7]. Usability and usefullness are investigated using feedback from radiologists on the quality of segmentation and on the learnability of the UI.

Elliott et al [8] conducted one of the first systematic clinical studies for comparing the usefulness of two interactive segmentation methods [9] embedded in a graphical user interface for radiation therapy planning. Segmentation was task-oriented and aimed at outlining fast the target volume and organs at risk for 3D radiation treatment planning. The strategy in [8] for integrating segmentation algorithms into a HCI for radiologists' use was to make sure that the user's clinical knowledge is efficiently complemented by the segmentation algorithm. According to [8], "…*automated image segmentation is in any case not wanted by the users. What they do want is a fast system (i.e. one that is faster than manual segmentation) in which the user has complete control over the results.*" Moreover, since radiologists tend to 'think in slices', 2D user interaction on a slice-by-slice basis was preferred over 3D interactive segmentation.

Shifting from slice-based 2D to full 3D user interaction is still an open question for diagnosis-oriented interfaces [10], which impacts not only on computer graphics, but also on computer vision techniques. Possible solutions for speeding up this shift focus on establishing new visualization standards for 3D image interpretation by radiologists [11], and on integrating 3D in the medical training curriculum [12].

Both 2D and 3D visualization paradigms were considered in the design of the 3D Slicer software and visualization platform [13], which offers a variety of tools for slice editing and interactive segmentation. The 3D Slicer interface is not intended for

clinical use, but its design considers a large diversity of user profiles including research-oriented physicians.

Research-oriented graphical interfaces. To date, there is a rich literature on algorithms for 2D and 3D medical image filtering, registration, segmentation and analysis [14] standing proof for the significant advances in image interpretation made in the last decade. To increase the visibility of such new methods, an extensive open-source library called the Insight Toolkit (ITK) has been built [15]. The algorithms available within ITK are programmed into C++ classes and may be considered as building blocks for a variety of task-oriented applications in medical imaging. While ITK is a valuable resource to the Computer Vision research community, it is of limited use to end-users such as radiologists and surgeons, since it does not provide a graphical interface for image visualization. The ANALYZE software system [16] is complementary to ITK, since it provides tools for interactive visualization. Augustine et al [17] proposed an integration of ITK and ANALYZE, which results in a user-friendly graphical interface for research-oriented end-users with little programming skills. The main modules of this tabbed interface offer various tools for filtering, registration and segmentation including control windows for parametric adjustment. A fast display updating mechanism provides visual information about the evolution of image processing algorithms. User error is minimized by constraining the sequential order of typical operations (i.e. noise reduction, edge detection, segmentation, analysis). Other integration efforts focused on adding dynamic visualization functionalities to ITK were reported by Rexilius et al [18] and Hansen et al [19].

Towards HCI design for collaborative and remote image analysis. Computer supported collaborative work is helpful in medical applications requiring the expertise of more than one physician or dedicated to clinical training. A task-oriented, collaborative interface for the visualization and analysis of fetal 3D ultrasound is proposed by Alberola-Lopez et al [20]. Their interface allows for session-based work, therefore minimizing errors caused by user fatigue. In [20], segmentation and/or 3D data manipulation are performed by one user who has acquired control through a token-grabbing paradigm; the other users are 'listeners' until the token is released. Text-based information exchange is asynchronuous, therefore feedback or discussions on a specific graphical model can take place at any time. The token-grabbing paradigm was also implemented in the Group-Slicer [21], the collaborative extension of 3D Slicer.

HCI for enhanced visualization during image-guided surgery. During open surgery, the surgeon has direct visual access only to exposed surfaces; the limitations are even more severe for minimally invasive surgery. The main consequence of limited surgical visualization is the non-accuracy of the pre-operative and intraoperative geometric localization of the targeted lesion. Three basic user requirements for the design of intraoperative image guidance interfaces are defined by McInerney and Roberts [22] as follows: a) give visual access to the structural lesion, b) enable the surgeon to define and verify the extent of resection, and c) facilitate the protection of normal healthy tissue.

While early image guidance systems used frame-based stereotaxy, frameless stereotactic systems provide tools for accurate navigation by relating the location of

instruments to preoperative, and more recently intraoperative image data. In the context of frameless stereotactic image guidance systems, computer vision techniques for image registration and segmentation are basic steps required for precise and interactive 3D rendering of the patient anatomy/physiology. An accurate stereotactic localization and digitization technique based on computer vision algorithms was reported in Heilbrun et al [23]. Their technique used a pair of 2D images acquired from different viewing angles with video cameras mounted on the ceiling of the operating room for determining the 3D location of markers seen in both images.

All stereotactic systems require co-registration, defined as a geometric mapping between at least two coordinate systems corresponding to the pre-operative and intraoperative spaces. A simple co-registration method requires known locations in both spaces of non-collinear points, defined either by fiducial markers or by natural landmarks. Computer vision algorithms for medical image registration also allow for rigid and non-rigid surface and contour matching. Pelizzari et al proposed in [24] a non-fiducial technique for registering MR, CT and PET brain images by matching the surface contour of the head. Anatomic surface curvatures may also be used for matching and registration purposes, as shown by Wang et al [25].

An interactive system for neurosurgery guidance and planning is presented in Gering et al [26]. Their system performs data fusion between various pre-operative and intraoperative scans by using 3D Slicer tools for image registration and segmentation. The system received positive clinical feedback for the graphical display of functional information, anatomical information, and information from contrast agents into a single view, as well as for the temporal data fusion. The main limitations were related to the surgeons' learning curve and to the use of the same interface for planning and for intra-operative guidance. The vast quantity of information extracted from data fusion and off-line analysis was found beneficial for planning, but overwhelming and distracting for on-line guidance.

In surgery guidance, image registration techniques play a central role; however, intraoperative image segmentation is likely to become a powerful tool in the process of image-guided interventions. Warfield et al [27] proposed a new intraoperative segmentation framework applied to the cryotherapy of liver cancer and to neurosurgery. Such a framework is designed to enable the monitoring of changes in anatomical structures (i.e. due to tumour resection) during surgery, and to quantitatively compare the progress of the interventional process with the preoperative plan.

3   Vision-Based Interfaces for the Operating Room

Computer Vision techniques such as markerless tracking of human motion and gesture recognition have been successfully integrated into perceptual user interfaces for applications such as video games, teleconferencing and surveillance. In particular, hand gesture recognition is useful for controlling the UI via command selection and virtual object manipulation. Surveys on hand gesture interpretation and on hand pose estimation are available in [28] and [29] respectively.

Operating rooms for minimally invasive surgery (MIS) are environments which could significantly benefit from using non-contact, gesture-controlled human-computer interfaces. Indeed, MIS procedures typically require computer support, and

standard computer peripherals are difficult to sterilize. Therefore, standard clinical protocols involve a human assistant who manipulates the computer display according to the surgeon's commands and needs for visualization. However, the assistant-in-the-loop approach is suboptimal and sometimes leads to frustration and prolonged time for performing the intervention.

The non-contact mouse proposed by Graetzel et al [30] enables the surgeon to directly control the user interface with simple hand gestures. The non-contact mouse supports the "wait-and-click" and "push-to-click" paradigms by hand tracking and gesture classification. The gestures of interest are simple and based only on the 3D position of the right hand; the hand motion is mapped to pointer movement using nonlinear gains, thus allowing for quick navigation and precise control. The non-contact mouse was successfully tested with a mock-up medical interface in the laboratory and in the operating room.

Face movement can also be used to control the user interface in MIS, as demonstrated by Nishikawa et al [31]. They proposed a face-tracking system that controls the laparoscopic camera positioning according to a face movement grammar. The tracker works with a 3DOF face pose by assuming that during the intervention, the surgeon's face remains almost parallel to and at a constant distance from the monitor screen. A user survey on a virtual testbed proved the usefulness of the system and its superiority over a voice-controlled interface; however, some tests suggested that face motion may distract the surgeon when performing very precise surgical actions.

While explicit user control over the graphical interface is absolutely necessary for the execution of critical tasks in laparoscopy, a certain degree of automation would speed up the intervention by allowing the surgeon to focus more on the surgery and less on the interface manipulation. As reported by Grange et al [32], process automation in a medical environment has to obey strict safety rules; thus, any automated user interface must be overridable by the surgeon's decision. The system proposed by Grange et al [32] combines gesture interpretation for explicit interaction with real-time monitoring of the surgeon's activity for automatically addressing some predictable surgeon's needs. They identify typical modes in a user interface for endoscopy, and prove that the transition from one mode to another can be automated using information from the visual tracking of the head, torso, and hands of the surgeon. The systems proposed in McKenna et al [33] and in Nishikawa et al [34] take a different automation approach by tracking instruments in the laparoscopic video instead of tracking the surgeon.

Video understanding techniques have also been developed for the quantitative assessment of basic surgical skills [35]. Such techniques may be good candidates for building multimodal HCIs which combine haptic [36] and visual information for evaluating laparoscopic and other surgical skills.

4   Perceptual Interfaces for Motor-Impaired Users

According to the Model Human Processor [37], a simple human computer interaction process comprises three cycles, namely perceptual, cognitive, and motor. Keates et al [38] proved that in standard UIs, motor impairment affects not only the motor cycle, but also introduces extra perception and cognitive cycles. To avoid extra cognitive

loads, the design of HCI for motor-impaired users must consider two alternatives: adapting the content of the graphical display and/or customizing input systems for allowing a more natural interaction. Computer Vision techniques such as real-time tracking of body features are suitable for implementing the second alternative.

In Morrison and Mckenna [39], the hand motion trajectory represents a basis for learning and recognizing hand gestures. Their system uses HMM models to learn and recognize a user-defined set of simple hand gestures which replace basic commands in a standard web browser (back, forward, start, open, OK, refresh, cancel, and close).

The system proposed by Betke et al [40] is able to track the motion of diverse body parts (nouse, eyes, chin, foot) with an algorithm based on spatiotemporal template matching. The body feature to be tracked can be specified by the user in the initialization phase. The motion of the tracked feature is then mapped onto the motion of the mouse pointer on the screen. The system in [40] proved to be useful for interaction based on dynamic or static item selection by pointing without clicking.

A binary selection paradigm in a visual UI can be controlled with eye-blinks or eye-brow motion, as shown in Grauman et al [41]. Applications based on eye-blink/eyebrow-raises input do not require mouse movement, since they are entirely controlled by clicks. Therefore, such applications implement a scanning mechanism which displays one option at a time until the user selects the desired option with a long eye-blink or with an eyebrow-raise. The eye-blinks and eyebrow-raises are detected by algorithms based on template matching and on the properties of eye motion during blinking.

While web browsers and educational games are useful tools for communication and learning, another basic need of motor-impaired users is related to moving in their physical environment. Yanco and Gips [42] proved that electrode-based gaze-tracking can be successfully integrated in the design of intelligent wheelchairs. In Kuno et al [43], the user can control the motion of an autonomous wheelchair via a perceptual interface which detects changes in his face direction using computer vision algorithms. This interaction paradigm is more natural than using the conventional joystick, since humans usually look in the direction they want to take. Visual information is collected with two video cameras, one observing the user and the other observing the environment. This visual information is seamlessly integrated with information from other types of sensors specific to autonomous vehicles in order to achieve the right balance between autonomous and user-defined motion. The robotic wheelchair is also able to observe the user at a distance, and to respond to the user's commands by recognizing hand gestures. This option is more suitable for elderly persons with limited capability of walking. More information on the design of vision-based systems for the assistance of elderly persons is to be found in the next section.
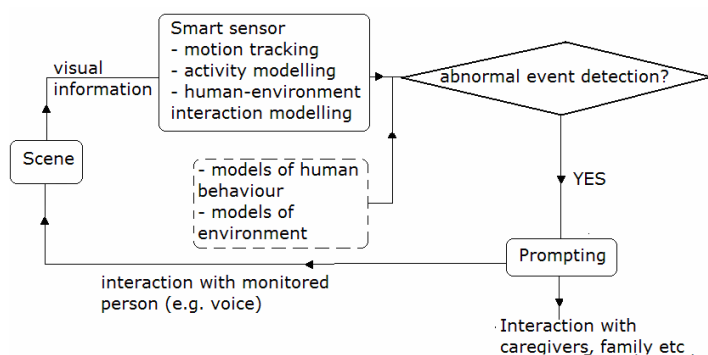
5   Vision-Based Intelligent Systems for Elderly Assistance

Pervasive computing is a promising technology for supporting aging-in-place. Indeed, intelligent environments can assist elderly persons in a supportive and non-intrusive way during their daily activities. Moreover, automatic visual monitoring may detect abnormal harmful events such as falls, loss of balance, or suspect periods of inactivity possibly caused by a stroke. Sensing agents based on computer vision are unobtrusive,

since they can be embedded in the environment without altering it. Vision algorithms are able not only to extract low-level data such as the subject's location and posture, but also to analyze human activities and interactions with the environment. Fig. 1 shows a generic vision-based monitoring system with its main modules, namely sensing, decision-making, and prompting.

However, the acceptance of vision-based monitoring in elderly health care is controversial, since it raises privacy and ethical concerns. User requirements for a fall detector in the context of a visual monitoring system were investigated by Mckenna et al [44]. They found that potential elderly users received well the idea of a vision-based monitoring system provided that images/videos are not stored or broadcast, and the visual input is analyzed only by a computer. Other findings in [44] were related to the design of the communication between the system and the faller (i.e. the prompting module in Fig. 1). Thus, the potential users wanted to be able to clear false alarms generated by the system, and also to press a button to call for help in case the fall has not been detected.

A significant segment of the community-dwelling elderly population suffers from various degrees of decline in cognitive functions and in memory. Such persons are unable to complete activities of daily living (ADL) [45] such as bathing and dressing on their own, since they do not remember the entire sequence of steps involved in the activity. A technique for visual ADL monitoring and assistance was proposed in Mihailidis et al [46]. This technique is applied to hand washing and integrates information from colour-based hand tracking and tracking of step-specific objects (i.e. soap bar). To date, computer vision techniques face limitations in fine motion tracking and ADL monitoring since they are very sensitive to contextual change. Pervasive computing systems such as the Aware Home project at Georgia Tech [47], the MIT's House_n project (http://architecture.mit.edu/house_n) and the Intel's Proact system [48] process information from various types of sensors for modelling ADLs. However, visual information plays an important role in activity recognition; indeed, [48] reported that hand washing is not well recognized by their system using radio-frequency identification tags, since water and metal absorb radio waves produced by these tags.

Fig. 1. Generic diagram of an intelligent system using a computer vision-based sensing agent

Visual monitoring of whole body motion finds relevant applications in systems designed to detect falls and unusual inactivity caused by stroke. The methods proposed by Nait-Charif and Mckenna [49] and by Sixsmith and Johnson [50] are based on a spatial model of the home environment composed of inactivity zones and entry (high traffic) zones. This model is learned in [49] from the spatio-temporal trajectories of the tracked subject and it can be used for fall and unusual activity detection. The approach in [50] is based on a user-defined spatial map, called risk map, and monitors subject with infrared cameras. Depending upon the zone type, their system adjusts the extent of acceptable zone inactivity. A subtle motion detector is correlated in [50] to the inactivity monitor in order to correctly detect some motion in static activities such as watching TV. Cuchiarra et al [51] proposed a video-based fall detector based on posture estimation and remotely connected with a PDA in order to enable an audio-video connection in case of emergency.

Whether focused on tracking fine motion for ADL assistance or on whole-body motion for fall or unusual inactivity detection, vision-based monitoring systems have to process data and prompt the users in real-time. A significant delay between the occurrence and the detection of an event is critical, since it propagates to the prompting module. Mihailidis et al [46] outlined that in ADL assistance, a delayed prompting for the completion of a particular step will result in user confusion. A late fall detection could have even more severe implications on the health and safety of the elderly.

6   Conclusions

In the context of user interface design, computer vision techniques play a dual role. First, computer vision algorithms for medical image understanding extract information from image data in order to provide assistance in processes such as diagnosis, therapy planning, and surgical navigation. *Such algorithms are designed to observe medical images*. As shown in Section 1, they may be successfully integrated in user interfaces for enhanced data visualization. Most important, differences in user requirements result in different algorithmic designs of computer vision techniques embedded in graphical interfaces, as well as in different interface design strategies. As an example, interactive and intuitive segmentation techniques in computer-aided diagnosis are preferred over automated techniques. Conversely, image-guided surgery uses 3D models of anatomical structures built from the off-line segmentation and multimodal registration of the raw image data; thus, segmentation and registration are not controlled by the user, and may be automated. The emphasis in image-guided surgery is on the accuracy and reliability of the patient-specific models built from preoperative data.

A second category of HCI-related computer vision techniques are designed to observe the user and/or understand his actions. Most often, action recognition is based on tracking the user's body parts. Techniques belonging to this second category are useful for controlling perceptual interfaces such as discussed in Sections 2 and 3. Moreover, human motion analysis can play an important role as input data in intelligent systems for monitoring the well-being of seniors. In the design of perceptual interfaces, the ease-of-use and the real-time response are essential user requirements. The real-time response also plays a critical role in monitoring human activities for elderly

assistance and support. As one may expect, other user requirements are task-oriented, such as addressing privacy and safety concerns in visual monitoring systems.

Computer vision techniques have already proven their usefulness for the design of medical research-oriented graphical user interfaces. It is expected that clinical acceptance will improve with advances of 'going filmless' in screening, diagnosis and therapy planning. The use of computers for inspecting images will probably trigger the use of interactive tools that augment visualization and improve the speed of diagnosis and planning.

Vision-based perceptual interfaces are currently among the latest trends in video games. As in the case of virtual reality applications, it is predictable that mature technologies and vision algorithms used in games will become transferable and/or adaptable to health applications such as described in Section 2 and 3. Multimodal interfaces integrating voice and visual recognition are also a promising alternative.

Vision-based monitoring for elderly assistance and support faces significant technological and ethical challenges. However, integrating video information with data gathered by other types of sensors, such as proposed in the design of intelligent environments can significantly improve the robustness of these systems. Respecting user requirements related to the privacy of video content, and promoting social connectedness via audio/video communications are viable strategies in coping with ethical concerns.

References

[1] J. Gosbee, E. Ritchie, "Human-computer interaction and medical software development." Interactions 4(4): 13-18, 1997.

[2] D. Pham, C. Xu, J. Prince, "Current methods in medical image segmentation." Annual Review of Biomedical Engineering, 2: 315-337, 2000.

[3] S.D. Olabarriaga , A.W.M. Smeulders, "Interaction in the segmentation of medical images: A survey." Medical Image Analysis 5: 127–142, 2001.

[4] G. Gerig, M. Jomier, M. Chakos, "Valmet: A new validation tool for assessing and improving 3D object segmentation." Proc. MICCAI Conf. Med. Image Comput. and Computer-Assisted Intervention, 516-523, 2001.

[5] S. Warfield, K.H. Zou, W. M. Wells, "Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation." IEEE Trans. on Med. Imag. 23 (7): 903-921, 2004.

[6] L. O'Donnell, C.-F. Westin, W.E.L. Grimson, et al, "Phase-based user-steered image segmentation", Proc. MICCAI Conf. Medical Image Computing and Computer-Assisted Intervention, 1022–1030, 2001.

[7] W. A. Barrett, E. N. Mortensen, "Interactive live-wire boundary extraction." Med. Image Anal., 1(4):331–341, 1997.

[8] P. J. Elliott, J. Diedrichsen, K. J. Goodson, R. Riste-Smith, G. J. Sivewright, "An object-oriented system for 3D medical image analysis." IBM Systems Journal, 35 (1): 4 – 24, 1996.

[9] P.J.Elliot, J.M. Knapman, W. Schlegel, "Interacting segmentation for radiation treatment planning." IBM Systems Journal, 31 (4): 620–634, 1992.

[10] M. Meissner, K. Zuiderveld (organizers), G. Harris, J. R. Lesser, A. Persson, M. Vannier (panelists), "End Users' Perspectives on Volume Rendering in Medical Imaging: A job well done or not over yet?" Panel, IEEE Visualization, Vis 05, Minneapolis, USA Oct. 2005.

[11] R. Shahidi, L. Clarke, R.D. Bucholz, H. Fuchs, R. Kikinis, R.A. Robb, M. Vannier, "White paper: challenges and opportunities in computer-assisted intervention." Comp. Aided Surgery, 6(3): 176-181, 2001.

[12] M. Meissner, B. Lorensen, K. Zuiderveld, V. Simha, R. Wegenkittl, "Volume Rendering in Medical Applications: We've got pretty images, what's left to do?" Panel, IEEE Visualization, Vis 02, Boston, USA Oct. 27-Nov. 1, 2002.

[13] D. T. Gering, A. Nabavi, R. Kikinis, et al, "An Integrated Visualization System for Surgical Planning and Guidance Using Image Fusion and Interventional Imaging." Proc. MICCAI Conf. Medical Image Computing and Computer-Assisted Intervention, 809–819, 1999.

[14] R. Robb, *Biomedical Imaging, Visualization and Analysis*, John Wiley and Sons, Inc: New York, 1999.

[15] L. Ibanez, W. Schroeder, L. Ng, J. Cates, *The ITK Software Guide*, 2003.

[16] D. Hanson, R. Robb, et al, "New software toolkits for comprehensive visualization and analysis of 3D multimodal biomedical images." Journal of Digital Imaging. 10(2), 1-2, 1997.

[17] K. Augustine, D. Holmes, R. Robb, " ITK and Analyze: A synergistic integration". Proc. SPIE Medical Imaging. (2004) pp. 6–15.

[18] J. Rexilius , W. Spindler, J. Jomier, et al," A Framework for Algorithm Evaluation and Clinical Application Prototyping using ITK." MICCAI Workshop on Open-Source Software 2005.

[19] N. Hanssen, B. von Rymon-Lipinski, T. Jansen, et al "Integrating the Insight Toolkit *itk* into a Medical Software Framework." Proc. of CARS Computer Assisted Radiology and Surgery 2002.

[20] C. Alberola-Lopez, R. Cardenes, M. Martin et al, "diSNei: A collaborative environment for medical images analysis and visualization". Proc. of MICCAI Medical image computing and computer-assisted interventions. pp. 814–23, 2000.

[21] F. Simmross-Wattenberg, N. Carranza-Herrezuelo, C. Palacios-Camarero et al, "Groupslicer: a collaborative extension of the 3D-slicer." Journal of Biomed. Informatics, 38: 431-442, 2005.

[22] J. McInerney, D.W. Roberts, "Frameless Stereotaxy of the Brain". The Mount Sinai Journal of Medicine, 67(4):300-310, 2000.

[23] M.P. Heilbrun, P. McDonald, C.Wicker et al, "Stereotactic localization and guidance using a machine vision technique". Stereotactic Functional Neurosurgery, 58:94-98, 1992.

[24] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring et al, "Accurate three-dimensional registration of CT, PET, and MR images of the brain", Journal Comput. Assist. Tomography, 13:20-26, 1989.

[25] Y. Wang, B.S. Peterson, L.H. Staib, 3D Brain Surface matching based on geodesics and local geometry, Computer Vision Image Understanding, 89:252-271, 2003.

[26] D.T. Gering, A. Nabavi, R. Kikinis et al, "An Integrated Visualization System for Surgical Planning and Guidance Using Image Fusion and an Open MR". Journal of Magnetic Resonance Imaging, 13:967–975, 2001.

[27] S. Warfield, A. Nabavi, T. Butz et al, "Intraoperative segmentation and non-rigid registration for image-guided therapy". Proc. of MICCAI, Medical Image Computing and Computer-Assisted Intervention, 176-185, Oct. 2000.

[28] V.I. Pavlovic, R. Sharma, T.S. Huang. "Visual interpretation of hand gestures for human computer interaction: A review," IEEE Trans. on Patt. Anal. and Machine Intelligence, 19 (7), 1997.

[29] A. Erol, G. Bebis, M. Nicolescu, R. Boyle, X. Twombly, "A Review on Vision-Based Full DOF Hand Motion Estimation", Proc. of the IEEE Workshop on Vision for Human-Computer Interaction (V4HCI), San Diego, Ca, June 2005.

[30] C. Graetzel, T.Fong, S.Grange, C.Baur, "A non-contact mouse for surgeon-computer interaction". Technology and Health Care, 12(3), 2004.

[31] A. Nishikawa, T. Hosoi, K. Koara et al, "Face mouse: a novel human-machine interface for controlling the position of a laparoscope", IEEE Trans. on Robotics and Automation, 19(5):825-844, 2003.

[32] S. Grange, T. Fong, C. Baur, "M/ORIS: A medical/operating room interaction system". Proc. of the ACM Int. Conf. on Multimodal Interfaces, State College, PA, 2004.

[33] S.J. McKenna, H. Nait Charif, T. Frank, "Towards Video Understanding of Laparoscopic Surgery: Instrument Tracking", Proc. of Image and Vision Computing, New Zealand, 2005.

[34] A. Nishikawa, S. Asano, R. Fujita et al "Robust visual tracking of multiple surgical instruments for laparoscopic surgery," Proc. of Comp. Assisted Radiology and Surgery, London, 2003.

[35] J. Chen, M. Yeasin, R.Sharma, "Visual modeling and evaluation of surgical skill". Pattern Analysis and Applications, 6:1-11, 2003.

[36] J. Rosen, M. Solazzo, B. Hannaford, M. Sinanan, "Objective Evaluation of Laparoscopic Skills Based on Haptic Information and Tool/Tissue Interactions". Computer Aided Surgery, 7(1): 49-61, 2002.

[37] S.K. Card, T.P. Moran, A. Newell, *The Psychology of Human-Computer Interaction*, 1983 Hillsdale, NJ: Lawrence Erlbaum Associates.

[38] S. Keates, P.J. Clarkson, P. Robinson, "Developing a methodology for the design of accessible interfaces." Proc. of the 4th Workshop on User Interfaces for All, Stockholm, Sweden, 1998.

[39] H. Morrison and S. J. McKenna, "Contact-free recognition of user-defined gestures as a means of computer access for the physically disabled". Proc. 1st Workshop on Univ. Access and Assistive Technology, Cambridge, UK, 99–103, Mar. 2002.

[40] M. Betke, J. Gips, P. Flemming, "The Camera Mouse: Visual tracking of body features to provide computer access for people with severe disabilities", IEEE Trans. on neural systems and rehabilitation eng., 10(1): 1-9, 2002.

[41] K. Grauman, M. Betke, J. Lombardi, J. Gips, G.R. Bradski, "Communication via eye-blinks and eye-brow raises: video-based human-computer interfaces", Univ. Access. Inf. Soc., 2: 359-373, 2003.

[42] H.A. Yanco, J. Gips, "Preliminary investigation of a semi-autonomous robotic wheelchair directed through electrodes," Proc. Rehab. Eng. Soc. of North Am.Annual Conf., 414-416, 1997.

[43] Y. Kuno, N. Shimada, Y. Shirai, "Look where you're going: A robotic wheelchair based on the integration of human and environmental observations." IEEE Robotics and Automation Magazine, 27-34, March 2003.

[44] S. J. McKenna, F. Marquis-Faulkes, P. Gregor, A. F. Newell, "Scenario-based drama as a tool for investigating user requirements with application to home monitoring for elderly people," Proc. of HCI Int., Crete, Greece, June 2003.

[45] S. Katz, "Assessing Self-Maintenance: Activities of Daily Living, Mobility, and Instrumental Activities of Daily Living." J. Am. Geriatrics Soc., 31(12): 721–726, 1983.

[46] A. Mihailidis, B. Carmichael, J. Boger, "The use of computer vision to support aging-in-place, safety, and independence in the home." IEEE Trans. on Inf. Tech. in Biomed., 8(3):238-247, 2004.

[47] Abowd, G. A. Bobick, I. Essa, E. Mynatt, W. Rogers, "The aware home: developing technologies for successful aging.", AAAI technical report.

[48] M. Philipose, K. Fishkin, M. Perkowitz et al, "Inferring activities from interactions with objects." IEEE Pervasive Computing, 3(4):50-57, 2004.

[49] H. Nait-Charif, S. J. Mckenna, "Activity summarization and fall detection in a supportive home environment", IEEE Int. Conf. on Pattern Recognition, 323-326, 2004.

[50] A. Sixsmith, N. Johnson, "A smart sensor to detect the falls in the elderly", IEEE Pervasive Computing, 3(2):42-47, 2004.

[51] R. Cucchiara, C. Grana, A. Prati, R. Vezzani, "Computer vision techniques for PDA accessibility of in-house video surveillance", Proc. of ACM Int. Workshop on Visual Surveillance IWVS, Berkeley (CA), 87-97, 2003.