

## TOOLSET 3D POSITION TRACKING FOR A VISIO-HAPTIC MIXED REALITY SYSTEM

**Mehmet Murat AYGÜN,  
Mehmet Ali GÜLER,  
Eren ÇELİK**  
Ankara, TURKEY,  
HAVELSAN Inc.

maygun@havelsan.com.tr  
mmaygun@etu.edu.tr  
mguler@havelsan.com.tr  
ecelik@havelsan.com.tr

**Hulusi BAYSAL, Ph.D.,  
Hacı Mehmet YÜKSEL  
Mehmet Akif NACAR, Ph.D.**  
Ankara, TURKEY,  
HAVELSAN Inc.

hbaysal@havelsan.com.tr  
myuksel@havelsan.com.tr  
manacar@havelsan.com.tr

**Yiğit TAŞCIOĞLU, Ph.D.**

Ankara, TURKEY,  
TOBB University of Economics  
and Technology

ytascioglu@etu.edu.tr

### ABSTRACT

Mixed reality simulation combines real and virtual objects in the same simulation environment. Applications that also include some form of haptic interaction are categorized as Visio-Haptic mixed reality (VHMR). With the advances in augmented reality technologies, there has been a growing interest on VHMR. Positive contribution of haptic interaction, especially in training simulators, is brought to attention in recent literature.

A haptic interface is designed to generate force feedback in a VHMR environment. This haptic interface allows trainees to use their real-life tools (i.e. scalpel, screwdriver). Real-life tools are attached to end-effector of the haptic system when haptic interaction is needed according to the shape of the virtual object. Unless a delicate position tracking system exist in VHMR simulations, it is impossible to evoke realism onto trainees. Therefore, to increase the quality of the realistic character of the simulations for VHMR systems, a 3D position tracking of the tool should be determined precisely and accurately.

This study focusses on precise, accurate and delicate enough 3D position tracking of slender real-life tools in a novel VHMR environment, where the free end of the tool is automatically engaged/disengaged by a haptic interface according to the haptic rendering requirements of the simulation. It is a fact that making interaction with real-life tools provides the trainee with a more realistic simulation experience if and only if a precise and accurate position tracking of those tools is calculated and fed into the system. A depth-camera based position tracking method is developed, marker-based and markerless tracking approaches are compared in terms of speed and precision in this study.

### ABOUT THE AUTHORS

**Mehmet Murat AYGÜN** is a Project Technical Manager at HAVELSAN Inc., where Mehmet Murat leads a civil aviation simulation development project. He worked more than five years in Defense Industry. During this period, he worked as control engineer, system engineer and project technical manager. He has experience in fixed/rotary wing aircraft flight simulators, Hardware in the Loop simulators, Haptics, system modeling & simulation and virtual/augmented reality technologies. Mehmet Murat is currently working towards the PhD degree in the field of visuo-haptic augmented reality at TOBB University of Economics and Technology (ETU). His on-going PhD thesis is focused on high fidelity Visio-haptic applications in medical simulation.

**Mehmet Ali GÜLER** received his B.S. (2017) in computer engineering from Gazi University. He worked at Infotron Inc. from August 2017 to December 2018. Since December 2018, he is working at HAVELSAN Inc. as Software Engineer.

**Eren ÇELİK** received his bachelor's degree in computer engineering from Gazi University in 2018. Since 2018, he is working on the Training and the Simulation Technologies Department at HAVELSAN Inc.

**Hulusi BAYSAL, Ph.D.** is Modeling and Software Group Manager at HAVELSAN Inc., where Hulusi leads a software developer team which is responsible for developing simulation software for training systems. He worked more than 11 years in Defense Industry. During this period, he worked as software engineer, team leader and software group manager. He has experience in fixed/rotary wing aircraft flight simulators, vehicle simulators, visual systems and virtual/augmented reality technologies.

Hulusi worked as an academic researcher between 2001 and 2008 in Dokuz Eylül University, Izmir. His major research focus was on medical simulation and 3D Visualization. He obtained MSc degree in 2004 and his Ph.D. degree in 2011 in computer science from Dokuz Eylül University.

**Hacı Mehmet YÜKSEL** received his B.S. (1999) in Computer Engineering and Information Science and M.Sc. (2002) in Computer Engineering from Bilkent University. Since October 2004, he has been working at HAVELSAN in various positions such as Software Engineer, IPT leader, Technical Manager, Group Manager and Project Manager.

**Mehmet Akif NACAR, Ph.D.** received his B.S. (1995) in Computer Engineering from TRAKYA University, M.Sc. (1998) in Electronics and Computer Engineering from GAZİ University, M.Sc. (2000) in Computer Science from Syracuse University and Ph.D. (2008) in Computer Science from Indiana University. He worked as a Department Chair between 2009 and 2014 at Harran University. He has extensive experience in cloud computing, grid computing, and virtual reality technologies. Since 2017, he is Executive Vice President of Training and Simulation Technologies at HAVELSAN.

**Yiğit TAŞCIOĞLU, Ph.D.** received his B.S. (2000) in Mechanical Engineering from Middle East Technical University, M.Sc. (2001) and Ph.D. (2006) in Mechatronics from Loughborough University. Since 2006, he is with the Mechanical Engineering Department, TOBB University of Economics and Technology. He is also the founding coordinator of the Mechatronics Minor Program.

## TOOLSET 3D POSITION TRACKING FOR A VISIO-HAPTIC MIXED REALITY SYSTEM

**Mehmet Murat AYGÜN,  
Mehmet Ali GÜLER,  
Eren ÇELİK  
Ankara, TURKEY,  
HAVELSAN Inc.**

**maygun@havelsan.com.tr  
mmaygun@etu.edu.tr  
mguler@havelsan.com.tr  
ecelik@havelsan.com.tr**

**Hulusi BAYSAL, Ph.D.,  
Hacı Mehmet YÜKSEL  
Mehmet Akif NACAR, Ph.D.  
Ankara, TURKEY,  
HAVELSAN Inc.**

**hbaysal@havelsan.com.tr  
myuksel@havelsan.com.tr  
manacar@havelsan.com.tr**

**Yiğit TAŞCIOĞLU, Ph.D.**

**Ankara, TURKEY,  
TOBB University of Economics  
and Technology**

**ytascioglu@etu.edu.tr**

### BACKGROUND

#### Mixed Reality

Mixed reality (MR) is a combination of real and virtual worlds to create new environments and views where physical and digital objects exist together and interact with each other in real time (Bhakta et al. 2005). It is also a mix of reality and mixed reality that contains augmented reality and augmented virtuality through immersive technology (Milgram and Kishino 1994).

See-through mixed reality devices project stereo display to glasses instead of screens. Stereo displays are created from one display for the right and left eyes individually and appear as single display for human vision. Thus, when the screen is too close to the human eye, no visual impairment occurs. This enables the trainee to see mixed reality application scenes such as holograms.

Similar to virtual reality devices, mixed reality devices must also calculate the position of the trainee's head to move the rendering camera in a mixed reality application, while the trainee is moving in real environment and visualizing the application scene (Milgram et al. 1994). Inside out tracking and outside in tracking are used in virtual and mixed reality devices. Outside in tracking contains two or more cameras which scan the environment to determine the position of the device. Inside out tracking scans the real environment with a camera and a sensor attached to the mixed reality device. With simultaneous localization and mapping (SLAM) technology, mixed reality devices create a map of the environment and determine their own position. This data is useful for navigation, odometry and robotic mapping in virtual or augmented reality (Durrant-Whyte and Bailey 2006; Henderson and Feiner 2011, Cadena et al. 2016).

Some features of the commercially available mixed reality glasses are listed in Table 1. Except Meta 2, the devices are standalone with their own operating systems, while Meta 2 requires a computer to work. The trainees interact with virtual objects via controllers and/or hand gestures. If an application has sound effects, the trainee can hear sound effects, which come from the virtual objects. Some mixed reality devices have additional light sensors that detect light intensity and light color. Developers can use these light sensor data to increase the reality of virtual objects by adding light intensity and light color to the mixed reality scene in Real-time.

In mixed reality applications the trainees can see, hear, and interact with virtual objects. Inclusion of haptics to the application enriches the reality of simulation by allowing the trainees to also be able to touch and feel the virtual objects. Positive contribution of Visio-Haptic Mixed Reality (VHMR), especially in training simulators, is brought to attention in recent literature (Cosco et al. 2013).

**Table 1. Commercial mixed reality devices**

Device	Field of View	Standalone	Processors	Input Method
Microsoft HoloLens	30°x17.5°	Yes	Intel 32-bit TPM 2.0 support Microsoft Holographic Processing Unit 1.0	Hand Gestures & Controller
Microsoft HoloLens 2	43°x29°	Yes	Qualcomm Snapdragon 850 Microsoft Holographic Processing Unit 2.0	Hand Gestures
Magic Leap One	40°x30°	Yes	2 Denver 2.0 64-bit cores + 4 ARM Cortex A57 64-bit cores NVIDIA Pascal™, 256 CUDA cores	Hand Gestures & Controller
Meta 2	90°x50°	No	-	Hand Gestures

### Hand Tracking in Mixed Reality

Mixed reality devices such as HoloLens and Magic Leap One cannot successfully scan small objects in short distance for spatial mapping data. This creates a problem when the trainee's hand enters between the trainee and the virtual objects, resulting in the virtual scene to appear over the trainee's hand. Even when the virtual objects have greater depth, they are shown on the trainee's hand like nothing is blocking the virtual scene. HoloLens 2 and Meta 2 have advanced hand tracking features but still have deviations and latency in position. The hand tracking results are shown in the application as black colored virtual hand models (Meta 2 also shows arm model). Since black colored objects are invisible in mixed reality views, hands are not shown in the scene view. Hence, if any objects in mixed reality scenes have greater depth value than the hands, their view remains behind the hands.

Dedicated hand tracking devices like Leap Motion (see Figure 1) are built for exact hand position and gesture tracking (Guna et al. 2014) and its tracking results are better than the ones of mixed reality devices. Leap Motion contains an infrared stereo camera system and connects to a computer with a USB cable. Leap Motion Software processes the tracking data and publishes to the SDK. The recent version of Leap Motion SDK supports Unity, Unreal Game Engine and publishes JSON tracking data over the WebSocket protocol. Leap Motion tracking data has all the position, rotation and length values that a high accuracy 3D hand model requires.



**Figure 1. Leap Motion device (Motion n.d.)**

Leap motion software publishes tracking data in JSON format over WebSocket to let the trainee get tracking data over a web connection. As our project runs on a mobile device, the mixed reality system gets the tracking data from WebSocket to our unity project. JSON data received from WebSocket is converted by a wrapper that we have developed for unity to show 3D modelled hands by already existing Leap motion Unity SDK. Using this method, the Unity leap motion SDK became usable over network connections.

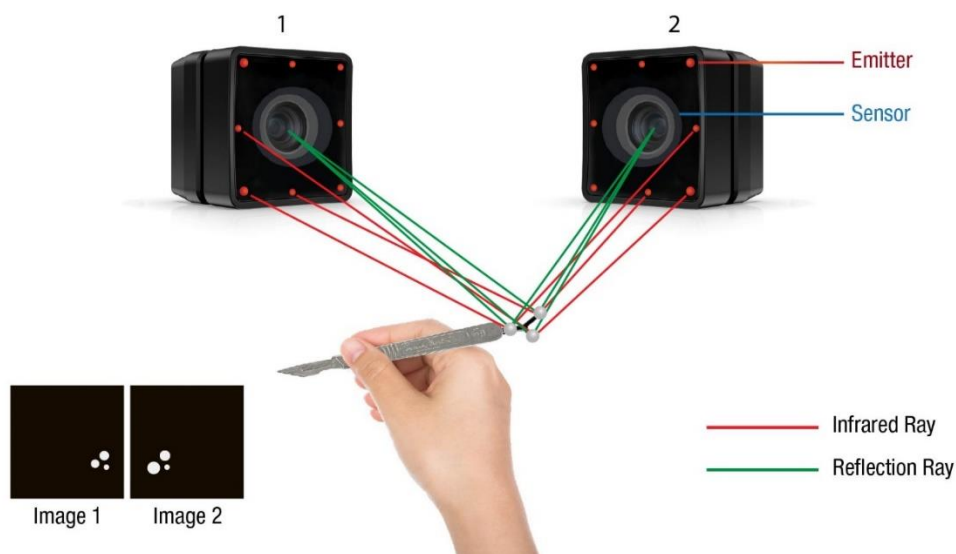
Created 3D modelled hands have the same size and position of trainee's hands. This method allowed us to transfer real hands to the virtual world and allowed trainees to interact with virtual objects in mixed reality applications. This 3D model also makes virtual objects stay behind the hands after it was colored black. Thus, hand tracking made virtual objects interactable with real hands and able to render virtual objects behind the hand.

### Optical 3D Object Tracking

This study focusses on 3D position tracking of slender real-life tools (i.e. scalpel, screwdriver) in a novel VHMR environment, in which the free end of the tool is automatically engaged/disengaged by a haptic interface according to the haptic rendering requirements of the simulation. It is a fact that making interaction with real-life tools provides the trainee with a more realistic simulation experience if and only if a precise and accurate position tracking of those tools is calculated and fed to the system. A depth-camera based method is developed, markered and markerless tracking approaches are compared in terms of speed and precision.

Positional tracking is detecting the position of objects (i.e. head-mounted displays, controllers or body parts) in Euclidean space in terms of rotation (pitch, yaw and roll) and translation data. Special sensors or markers may help to identify the objects to be tracked (Harders et al. 2009). Accurate position tracking is essential while working with real-life tools in a VHMR simulation environment. The Haptic system needs a precise and accurate position of real-life tool position to interact with when the tool and haptic engaged/disengaged according to the training scenario. With this way, more realistic interaction between trainees and VHMR system including virtual objects could be provided and this would eventually increase the quality of training simulations.

Optical object tracking is possible with a set of computer vision algorithms and tracking devices (i.e. camera, stereo camera depth camera). It has the same vision principle as stereoscopic human vision. Human vision is binocular vision and when humans look at an object, they can determine the approximate distance of the object. However, installing a pair of cameras is not enough to simulate human vision since cameras must be calibrated to get the location of the object. Humans calibrate their vision early in life by correlating the location of the object with their outstretched hands as an infant. Optical systems are affordable and reliable; however, they are also hard to calibrate since they require direct line of light without occlusions to receive correct data.



**Figure 2. Tracking a tool with spherical markers using two infrared cameras**

In markered object tracking, targets have their own markers which form a known pattern (Figure 2). Markers can be passive, like spheres or QR codes, or they can be active light sources. One or more cameras seek to get the position of markers with various algorithms. Tracking algorithms have to contend with missing data in the case where one or more markers are not being detected for any reason. Since blocking the unwanted infrared light in the tracking area is relatively easy, using infrared markers with synchronized cameras enables object tracking without much scattering (Mehling 2006). In a markerless approach, tracking cameras continuously search for 3D objects and compares them with known 3D models. This technology is used by Leap Motion and other mixed reality devices like Microsoft HoloLens and Meta 2. (Tagliasacchi et al. 2015). Some commercial tool tracking systems are listed in Table 2.

**Table 2. Commercial tool tracking systems**

Device	Max. Distance	Min. Distance	Tracking Accuracy	Camera Type	Marker
PST Pico	1.5 m	5 cm	<0.5 mm	Infrared	Yes
FusionTrack 250	2.4 m	400 mm	0.09-0.54 mm	Infrared	Yes
OptiTrack	12.2 m	-	<0.1 mm	Infrared	Yes
MicronTracker HX-40	120 cm	15 cm	0.20 mm	Infrared	Yes
Leap Motion	82.5 cm	31.7 cm	>1 cm	Monochrome	No
Structure Core	5 m	0.3 m	>1 mm	RGB	Yes/No

## Haptics

The term Haptic, originates from the Greek word “haptós” which means pertaining the sense of touch in Greek language. Haptics technology creates an artificial sense of touch by applying vibrations, motions or force (Morris et al. 2007). It is also called kinesthetic communication or 3D touch (Ott, Thalmann, and Vexo 2007). Some simple yet common haptic devices are game controllers, joysticks and steering wheels. More advanced haptic devices are employed especially in surgical training applications and high precision teleoperation.

Haptics technology allowed for the creation of controlled haptic virtual objects. Thus, investigation of how human sense of touch works became possible. There are three sensory systems in humans about the sense of touch; cutaneous, kinesthetic and haptic (Srinivasan and LaMotte 1995, Freyberger and Färber 2006). Cutaneous and kinesthetic are perceptions which mediate named tactual perception (Bergmann Tiest and Kappers 2009). Haptic sensory is an active touch to recognize or communicate with objects (Bergmann Tiest 2010).

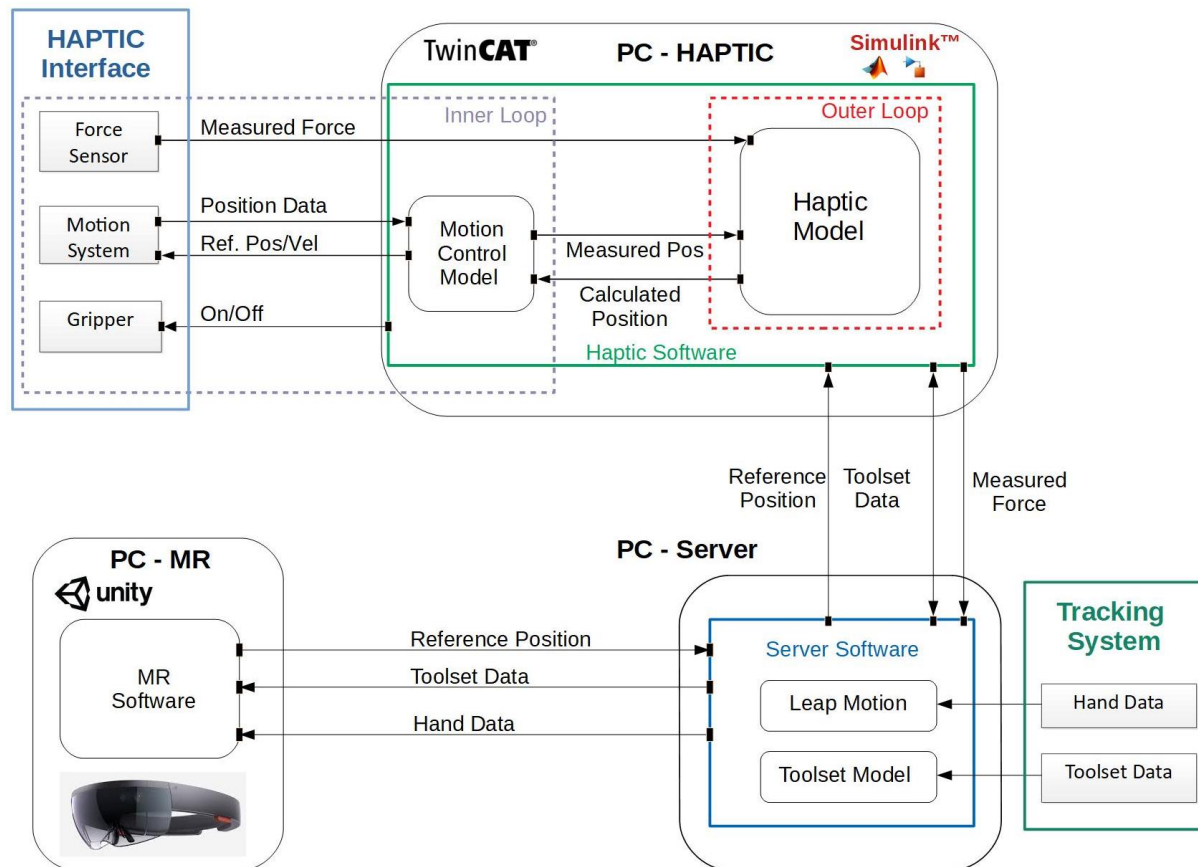
In most of the haptic-assisted simulation environments, the trainee has direct contact with the physical haptic device. Users are generally distracted by this situation, which reduces the realism of the simulation environments and the quality of training. Because instead of feeling the weight and sensation of the tool, they feel the physical haptic device itself.

## VISIO-HAPTIC MIXED REALITY SYSTEM

Mixed Reality (MR) typically deals with the visual addition of virtual objects to a real scene. However, to fully experience the mixed environment, the integration of virtual and real objects must be extended to the rest of the sensory modalities. VHMR (Visio-Haptic Mixed Reality) is a trainee who can see and touch dynamic virtual objects with static real objects on the stage (Cosco et al. 2013).

Figure 3 depicts the architecture of the developed VHMR system and Figure 4 shows the general layout. The system is comprised of a Microsoft HoloLens mixed reality device, a Leap Motion device for hand and tool tracking, a haptic interface, and two computers, one for haptics and other to act as a server. The server application creates communication interface between the other components.

Firstly, the trainee marks the reference position with the mixed reality device to make the entire system work in the same 3D space. Then, the virtual object to be visualized is selected and its mesh and material data is sent to the server application. In return, the mixed reality device collects hand and tool tracking data from Leap Motion, which is connected to the server computer, over the server application and displays if necessary.



**Figure 3. VHMR system architecture**

Tool tracking data contains the tool tip position and tool direction. Hand tracking system sends hand orientation data to VHMR system over server application. Hand tracking serves two purposes; first is to allow the trainee to interact with the virtual object by hand without using a specific tool, and the second is to render the virtual object behind the hand in case the hand enters between the trainee and the virtual object. When trainee interacts virtual objects with the tool, the haptic system obtains tool force that applied by the trainee and sends it to the VHMR system over server application to use it on virtual feedback effects.

Server application collects and transfers data to needed systems. The system collects hand and tool data from hand and tool tracking system. Reference position and mesh data that comes from the mixed reality system and measured force that comes from the haptic system are transferred by the server application.

The haptic system operates with reference position, mesh and material data that received from the mixed reality system. The system also gets tool data to follow tool before interaction. In robotic mode, the haptic system catches tool by grasping mechanism and track tool data for more accurate tracking and sends data to the mixed reality system. Before trainee interacts virtual objects with the tool, the haptic system gets tool tracking data to follow tool. When trainee interacts virtual object with the tool, the haptic system holds the tool with retention mechanism by checking the distance with an optical distance sensor. When the trainee moves tool, the haptic system in robotic mode also moves and follow the tool. Also if trainee applies power to a virtual object, the haptic system in haptic mode creates force feedback according to the virtual object's material. When the trainee pulls the tool back from the virtual object,

the interface switches to the robotic mode. The gripper releases the tool and follows the tool's point as closely as possible to catch it back when necessary.

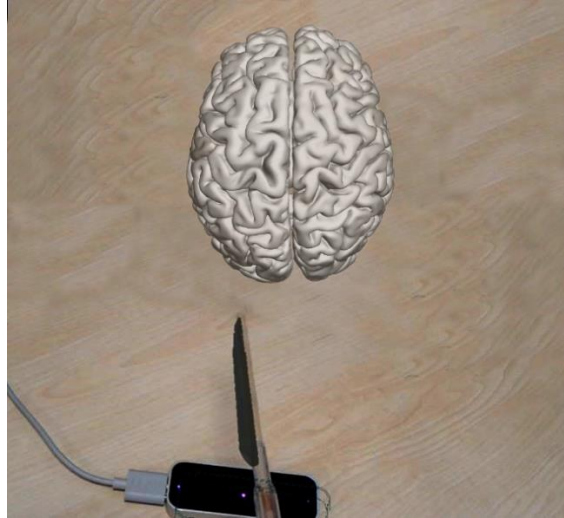


**Figure 4. General Design of Haptic System**

## **RESULTS AND FUTURE WORK**

This study is focused on 3D position tracking of real-life tools; device capabilities and tracking methods are examined. This paper reports the initial results of a work in progress.

Figure 5 shows a markerless tracking image of a slender object with Leap Motion and Microsoft HoloLens, as they appear to the trainee. A difference of up to 2 mm exists between the actual and virtual tool positions during translation. In the case of rotation, the difference becomes much larger. The targeted tool tracking accuracy of our application is 1 mm.



**Figure 5. Tool tracking with Leap Motion and Microsoft HoloLens**

According to infrared camera system specification, infrared cameras can track marker mounted object in 1 m<sup>3</sup> are with low latency and accuracy value below 1 mm. This method is usable for our tracking system in Visio-Haptic mixed reality system. Marker mounted objects enable any tool to be usable for our tracking system. Software update is not necessary in case of tool change.

The haptic interface needs a precise and accurate position of the real-life tool to engage/disengage according to the training scenario. Therefore, a more realistic interaction between trainees and VHMR system including virtual objects could be provided and this would eventually increase the quality of training simulations. When haptic perception requirements and the initial tracking results are considered, markerless tracking seems not to be a good choice. Hence, we are inclined towards marked tracking with infrared cameras.

## ACKNOWLEDGEMENTS

This study is financially supported by the Scientific and Technological Research Council of Turkey and HAVELSAN Inc. as a T1501 project under grant number: 3180230.

## REFERENCES

- Bergmann Tiest, W. M. and A. Kappers. 2009. "Cues for Haptic Perception of Compliance." *IEEE Transactions on Haptics* 2(4):189–99.
- Bergmann Tiest, Wouter M. 2010. "Tactual Perception of Material Properties." *Vision Research* 50(24):2775–82.
- Bhakta, Dee, Isabel dos Santos Silva, Craig Higgins, Leena Sevak, Tashmin Kassam-Khamis, Punam Mangtani, Herman Adlercreutz, and Anthony McMichael. 2005. "A Semiquantitative Food Frequency Questionnaire Is a Valid Indicator of the Usual Intake of Phytoestrogens by South Asian Women in the UK Relative to Multiple 24-h Dietary Recalls and Multiple Plasma Samples." *The Journal of Nutrition* 135(1):116–23.
- Cadena, Cesar, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, Jose Neira, Ian Reid, and John J. Leonard. 2016. "Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age." *IEEE Transactions on Robotics* 32(6):1309–32.
- Cosco, Francesco, Carlos Garre, Fabio Bruno, Maurizio Muzzupappa, and Miguel A. Otaduy. 2013. "Visuo-Haptic Mixed Reality with Unobstructed Tool-Hand Integration." *IEEE Transactions on Visualization and Computer Graphics* 19(1):159–72.
- Durrant-Whyte, Hugh and Tim Bailey. 2006. "Simultaneous Localization and Mapping: Part I." *IEEE Robotics and Automation Magazine* 13(2):99–108.
- Freyberger, Franziska K. B. and Berthold Färber. 2006. "Compliance Discrimination of Deformable Objects by Squeezing with One and Two Fingers." Pp. 271–276 in *Proc. EuroHaptics*. Vol. 2006.
- Guna, Jože, Grega Jakus, Matevž Pogačnik, Sašo Tomažič, and Jaka Sodnik. 2014. "An Analysis of the Precision and

- Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking.” *Sensors* 14(2):3702–20.
- Harders, M., G. Bianchi, B. Knoerlein, and G. Szekely. 2009. “Calibration, Registration, and Synchronization for High Precision Augmented Reality Haptics.” *IEEE Transactions on Visualization and Computer Graphics* 15(1):138–49.
- Henderson, Steven and Steven Feiner. 2011. “Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair.” *IEEE Transactions on Visualization and Computer Graphics* 17(10):1355–68.
- Mehling, Michael. 2006. “Implementation of a Low Cost Marker Based Infrared Optical Tracking System.” *Master Thesis* 101.
- Milgram, P., H. Takemura, a Utsumi, and F. Kishino. 1994. “Mixed Reality ( MR ) Reality-Virtuality ( RV ) Continuum.” *Systems Research* 2351(Telemanipulator and Telepresence Technologies):282–92.
- Milgram, Paul and Fumio Kishino. 1994. “Taxonomy of Mixed Reality Visual Displays.” *IEICE Transactions on Information and Systems* E77-D(12):1321–29.
- Morris, Dan, Tan Hong, Federico Barbagli, Timothy Chang, and Kenneth Salisbury. 2007. “Haptic Feedback Enhances Force Skill Learning.” *Proceedings - Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, World Haptics 2007* 21–26.
- Motion, Leap. n.d. “Leap Motion.” *Leap Motion*. Retrieved May 24, 2019 (www.leapmotion.com).
- Ott, Renaud, Daniel Thalmann, and Frédéric Vexo. 2007. “Haptic Feedback in Mixed-Reality Environment.” *Visual Computer* 23(9–11):843–49.
- Srinivasan, M. A. and R. H. LaMotte. 1995. “Tactual Discrimination of Softness.” *Journal of Neurophysiology* 73(1):88–101.
- Tagliasacchi, Andrea, Matthias Schröder, Anastasia Tkach, Sofien Bouaziz, Mario Botsch, and Mark Pauly. 2015. “Robust Articulated-ICP for Real-Time Hand Tracking.” *Computer Graphics Forum* 34(5):101–14.