A design for a mobile robotic avatar - modular framework

Seet, Gerald Gim Lee; Pang, Wee Ching; Burhan; Chen, I-Ming; Iastrebov, Viatcheslav; Gu, William; Wong, Choon Yue

2012

Seet, G. G. L., Pang W. C., Burhan, Chen, I.-M., Iastrebov, V., Gu, W., & Wong, C. Y. (2012). A design for a mobile robotic avatar - modular framework. 2012 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON).
doi:10.1109/3DTV.2012.6365444

<https://hdl.handle.net/10356/97973>

<https://doi.org/10.1109/3DTV.2012.6365444>

© 2012 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at:
<https://doi.org/10.1109/3DTV.2012.6365444>

Downloaded on 26 Apr 2023 18:33:22 SGT

A DESIGN FOR A MOBILE ROBOTIC AVATAR – MODULAR FRAMEWORK

Gerald Seet^{}, Pang Wee Ching[#], Burhan[#], Chen I-Ming^{*},
Viatcheslav Iastrebov^{*}, William Gu[#], Wong Choon Yue[#]*

School of Mechanical & Aerospace Engineering^{*}, and
BeingThere Centre[#], Institute for Media Innovation,
Nanyang Technological University, Singapore

ABSTRACT

This paper presents a modular architecture for MAVEN; a mobile robotic avatar. The proposed modular architecture allows for rapid customization to suit changes in application requirements. Aspects of modularity were applied to both hardware and software subsystems in the form of modules for mobility, head and arm gesturing, internet telephony, artificial intelligence (AI) power and communication. Module interconnectivity is achieved using industry standard Ethernet and serial communication standards. This allows modules the flexibility of implementing different operating systems and leveraging on existing development.

By grouping hardware and its complementary software into ‘self-contained’ modules, changes in requirements can be implemented with relative ease; configurations can vary from an intelligent information kiosk to an autonomous mobile Avatar, with arms and head units that permit gesturing. The modular approach allows for the reuse of modules, enhancing ease of reconfiguration and cost savings.

Index Terms — Avatar, System Architecture, Modular Architecture, Modular Design, Robotic Platform

1. INTRODUCTION

MAVEN is an acronym for “Mobile Avatar for Virtual Engagement by NTU”. It is an intelligent mobile telepresence system that permits the user (“Inhabitor”) to attend an overseas conferences or meeting without leaving the comfort of one’s office. MAVEN, as shown in Figure 1, is a synergistic fusion of intelligent mobile robot technology enabled with teleconferencing and telepresence functionality[1]. It allows its user to “move” around on a mobile platform, at walking speed and establish one’s presence with the use of a 2D graphical, or 3D physical, display. The audio and video link permits the transmission of speech and images from the inhabitor’s site to the place where MAVEN resides.

Telepresence [2] may enable a remote human to effectively interact through the provision of all the sensations of actually being at the remote site. The need for mobility seriously increases the cognitive load of the “inhabitor” and distracts the user from the task of interaction. Under normal conditions, humans move about freely with little or no conscious cognitive effort.

Current mobile telepresence systems [3][4] have limited motion autonomy [5]. Its controls require a high level of dexterity and imposes significant cognitive burden on the human “Inhabitor”. In an effort to reduce the cognitive load resulting from motion related activity, semi-autonomous motion navigation [6] is required to perform behaviours like obstacle and collision avoidance, human following as well as other human motion

behaviours. To encourage the interaction between humans and to enhance the ease of use, the avatar must possess human like capabilities and behaviours. Such capabilities will allow the “inhabitor” to concentrate on the primary task of interaction with other humans.

Studies indicate that up to 93% of communication effectiveness is determined by nonverbal cues. Another study indicated that the impact of a drama performance attributes 45% to speech and 55% to nonverbal communication [7]. To support additional communication modalities, MAVEN will be enabled with the capability to convey head, face and arm gestures. The avatar can project the inhabitor’s face textures on its robotic head and mimic her head movements via a rear projection robotic head with pan-tilt mechanism to convey head movements. Light weight arms and hands are also designed to enable the portrayal of arm and hand gestures.

Figure 1. MAVEN Platform

Challenges to the realisation of a viable mobile avatar hinges on reducing data bandwidth, response latency, ease-of-use and enhancing operational safety. Safety is essential as both humans and the robot avatar share the same environment. The design approach is to transmit high level data, commands and to imbue the platform with autonomous behaviours, as a means of enhancing performance and to mitigate safety issues. In the absence of a data communication link, the avatar is expected to act autonomously in a safe and reliable manner.

Optimum system configurations are highly dependent on application requirements. The option of being able to easily reconfigure and replace different modules is an advantage. In order to accommodate the needs of different applications and their required functions, a suite of core and optional modules are proposed.

A brief description of basic functions and their corresponding modules are listed:

- **Power Supply** - This module provides regulated direct current (DC) supply at a range of voltages; 5V DC, 12V DC etc.
- **Host Subsystem** - This is the main processor module which synchronises the data and command flow among all the peripheral modules. The operating system is Windows 7 but can have options for Linux or other operating systems.
- **Mobility** – Essentially, this is the physical mobile platform, motion controller and suite of basic ranging and collision sensors. In addition, the module includes algorithms for the basic motion behaviours like map-based navigation, way-point navigation, obstacle avoidance and collision prevention.
- **Application Sensor Modules** - Depending on the nature of application and the intended deployment, the sensors could include the Kinect sensor, light-detection-and-ranging (LIDAR) scanners, webcams and microphones.
- **Communication** - This defines the wireless RF Ethernet connectivity to support Internet telephony and data transfer to a remote operator (inhabitor) console.
- **Basic Screen Display** - A flat screen display is invariably, the primarily media for displaying information. At the remote site, text and graphical information can be displayed.
- **Remote Inhabitor Console** - This provides the human inhabitator with telepresence empowered interface for controlling the remote mobile platform and the capability to support interaction with other persons and systems at the remote site. The station incorporates a host computer to support human-computer interactions (HCI). Much like the work described in [8], here, user interface design for the console has also considered the use of both the visual as well as the auditory sensory channels in presenting information to the console's user or inhabitator. This emphasis on enhancing immersiveness is due to the fact that telepresence is an essential element of teleoperated systems [9].
- **Internet Telephony** - Primarily this is the backbone of the video and audio interconnectivity between the inhabitator and remote console. This module connects through the communication module.
- **Head Module** - This is a 3D robotic head module capable of projecting a textured image of the inhabitator. In addition a two DOF neck mechanism allows head gestures to be conveyed as in a nod or pan motion.
- **Arm Module** - This provides the ability to convey basic hand and arm gestures. Later enhancements could include the ability for physical interaction like; handshake and passing of small objects.
- **AI and Application Behaviour** - This would invariably need to be customized, depending on the application. There are however common preprocessing functions that can be generic and reconfigurable.

Depending on the potential requirements, the different modules are selected and integrated. By adopting a modular approach, "standard" modules may be created for a variety of different applications.

2. SYSTEM & MODULE CONFIGURATION

Aspects of modularity were applied to both hardware and software subsystems in the form of modules for mobility, head and

arm gesturing, internet telephony, artificial intelligence (AI) power and communication. The Ethernet bus was chosen for its high bandwidth and popularity, which brings with it economic benefits of lower component cost. As Ethernet interconnection is support by all major operating systems and programming languages, this allows modules the flexibility of implementing different operating systems and leverages on existing development.

Whilst Ethernet connection may be amply supported by the RJ45 connector, power connection was generally inadequate over the PoE (Power over Ethernet) standard. Electrical supply is provided over separate electrical connection of 5V DC or 12V DC.

2.1. System Architecture

Figure 2. System Architecture

Figure 2 shows the basic system architecture. The inhabitator console communicates with the remote avatar via the Internet network. Through the Internet there are two TCP/IP and one UDP/IP streams which provide command, video and audio services. To enhance data bandwidth utilisation, audio and video data are encoded before transmission and decoded on receipt of transmission. The inhabitator console may take the form of a simple flat display and keyboard, or enhanced with joysticks, pedals, immersive stereo displays [10] and motion trackers.

2.2. Module Architecture

Figure 3. Basic Module Configuration

The basic module comprise of a communication and supply sub modules, with an embedded processor or computer. The modules are connected to the host sub system module via an RJ-45 Ethernet connection and an external DC power source. To support module functionality, there will invariably be sensor and actuator options. The embedded processor includes the necessary software algorithms for the processing of data and control of actuators. This architecture is depicted in the Figure 3, above.

Taking the head module as example, the embedded processor is an embedded computer or micro-processor. The actuator sub module would comprise of the pico-projector and neck actuators. In the initial implementation, there are no external sensors.

3. IMPLEMENTATION – MAVEN PLATFORM

Figure 4 shows the detail architecture of MAVEN. The platform connects to an inhabitator console via the Internet, as illustrated in Figure 2, which depicts MAVEN and her remote operator console.

Figure 4. Architecture of MAVEN

MAVEN incorporates the mobility module which imbues the platform with autonomous navigation and holonomic motion capabilities. The holonomic platform includes mobility sensors like the Hokuyo LIDAR scanner and infrared collision sensors. The navigation algorithms for obstacle avoidance and collision detection are executed on the Host Subsystem.

The flat screen display projects the image of a virtual avatar or that of the human inhabitator at the remote operator console. Through the Internet telephony module, the inhabitator can interact with the avatar's environment, seeing a wide 170° field-of-view (FOV) image of the remote scene.

Three interaction modes were implemented for the human inhabitator to control the motion of the mobile platform. The operator can “move around” and communicate with others at the remote site using the avatar. The inhabitator can control MAVEN's camera and drive the robot around by issuing motion commands either from an attached joystick or via clicking a point on a map.

It is evident that autonomous navigation is highly beneficial in reducing the cognitive demands required in remotely operating a robot [9][10]. Imbuing MAVEN with the capability to navigate autonomously is only a beginning to the provision of intelligent autonomous navigation of a “being there” avatar. The telepresence enabled robot should also maintain a comfortable and safe distance from people who are co-located with the robot. This facilitates interaction in addition to safe and effective autonomous navigation. In creating a robot for interacting with other humans, it must be equipped with human like motion attributes to promote a safe, natural and comfortable interaction.

Future interaction options include facial and hand gesture interaction capabilities to further enhance the quality of interaction for the user.

4. OTHER POSSIBILITIES

Numerous other configurations are possible as in the case of an intelligent information kiosk. A standalone system would need the minimum of modules. This would include:

- Power Supply
- Host Subsystem
- Basic Screen Display

- AI & Application Behaviour
- Application Sensor Modules
- Head Module
- Arm Module

Figure 5 shows a standalone static configuration in terms of its functional modules and physical implementation. Adding mobility as in the application of a museum robotic guide would require additional modules. The task of an automated museum guide is generally structured. This could be implemented as a sequence of audio/video narratives and platform translation movements, repeated from one exhibit to the next exhibit. The possibility of unstructured intervention by way of unexpected questions and other contingencies (collisions) could be addressed by direct human intervention when the “robot” is unable to proceed within a structured sequence.

Figure 5. Static Standalone Configuration

With the need of a human supervisor (inhabitor) and mobility, additional modules are required, including:

- Mobility
- Communication
- Remote Inhabitor Console

The mobility module would permit the mobile avatar to be directly steered. The communication module will support visual and audio transmission between the inhabitator station and the avatar. The inhabitator station provides the opportunity for remote observation, monitoring and intervention, as required. With the desire for enhanced interaction, additional modules may be incorporated, including:

- Internet Telephony
- Head Module
- Arm Module

This will allow direct face-to-face interaction between remote parties with head, arm and hand gestures with accurate remote facial representation of the human supervisor.

Figure 6. Head Module

The head module is, geometrically, an anthropometrically accurate representation of a human head. It has a neutral male face contour. A face texture can be rear projected to represent any specific individuals. During a telepresence session, it can project the face of the inhabitator.

The entire system comprises of three main components; a matt translucent face, a pico-projection system, as well as a pan-tilt mechanism. This is housed within the space of a human head. The entire projection module is encapsulated within the head, and has no protruding components. This is novel in that all other similar systems have an open back [11]. The compact design is achieved by the use of a custom lens and mirror projection system [12].

Figure 7. Arm Module

The arm module is similar in architecture to the head module. It adopts a similar Ethernet and power interconnection. The module is designed to receive both low-level motion commands and high level gesture commands like a hand wave, arm raised or lowering of the hands. Extension of capabilities will enable a handshake motion or interaction with other objects using a hand with articulating fingers.

The upper and lower arms are constructed from laser cut corrugated paper material and reinforced with light weight plywood. To simplify implementation, Dynamixel[®] robot servo motors were used to drive the shoulder and elbow joints. Such actuators are termed intelligent motors with integrated controller, driver and network functionality. It has the ability to maintain position with speed control.

There are two degrees-of-freedom (DOF) at the shoulder joint and two more at the elbow. This configuration with reduced DOF was deemed adequate for a limited range of arm and hand gestures.

5. DISCUSSION AND CONCLUSION

With a modular design there is invariably an adoption of distributed control architecture. This encourages the use of multi-processors and parallel processing to enhance system performance, and a natural segmentation of functionality, leading to easier programming and a reduced data bandwidth requirement. The modules have been applied to a number of different application configurations with relative ease. A secondary purpose of adopting this modular approach, and incorporating industry standard communication interface, is to realize a suite of commercial-off-the-shelf (COTS) modules to hasten future development of robotic avatar systems. In the near future, developers can focus on application specific software development without the need to be distracted by electro-mechanical module design and their physical interconnection.

ACKNOWLEDGEMENT

This research, which is carried out at the BeingThere Centre, is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

REFERENCES

- [1] G. G. L. Seet, W. C. Pang and Burhan, "Towards the Realization of MAVEN - Mobile Robotic Avatar," in *The 25th International Conference on Computer Animation and Social Agents*, Singapore, 2012.
- [2] M. Minsky, "Telepresence," *Omni*, vol. 2, pp. 45--52, 1980.
- [3] E. Guizzo, "When my avatar went to work," *Spectrum, IEEE*, vol. 47, pp. 26--50, 2010.
- [4] VGo Communications, Inc., May 2011. [Online]. Available: <http://www.vgocom.com>.
- [5] K. M. Tsui, M. Desai, H. A. Yanco and C. Uhlik, "Exploring use cases for telepresence robots," in *Proceedings of the 6th international conference on Human-Robot Interaction*, 2011.
- [6] K. W. Ong, G. Seet and S. K. Sim, "An implementation of seamless human-robot interaction for telerobotics," *International Journal of Advanced Robotic Systems*, vol. 5, pp. 167--176, 2008.
- [7] A. Mehrabian, *Nonverbal communication*, 3rd ed., Aldine Transaction, 2009.
- [8] C. Y. Wong, G. L. Seet, S. K. Sim and W. C. Pang, "A hierarchically structured collective of coordinating mobile robots supervised by a single human," in *Mobile Ad Hoc Robots and Wireless Robotic Systems: Design and Implementation*, IGI Global, 2012, To be published.
- [9] S. Tachi, H. Arai, I. Morimoto and G. Seet, "Study on Tele-existence – Mobile telerobot systems with sensation of presence," in *25th Annual Conference of the Society of Instrumentation and Control Engineers*, Tokyo, 1986.
- [10] S. Tachi, H. Arai, I. Morimoto and G. Seet, "Feasibility experiments on a mobile tele-existence system," in *The International Symposium and Exposition on Robots (19th ISIR)*, Sydney, Australia, 1988.
- [11] T. Yotsukura, S. Morishima, F. Nielsen, K. Binsted and C. Pinhanez, "HyperMask-- Projecting a Talking Head onto a Real Object," *The Visual Computer*, vol. 18, pp. 111--120, 2002.
- [12] W. C. Pang, Burhan and G. Seet, "Design Considerations of a Robotic Head for Telepresence Applications," in *International Conference on Mechanical Engineering and Mechatronics*, Montreal, Canada, 2012.