

# Interacción Humano-Computadora: Personajes Virtuales

N. Jofré, G. Rodríguez, Y. Alvarado, J. Fernández, R. Guerrero

*LCG - Universidad Nacional de San Luis*

*Ejército de los Andes 950*

*Tel: 0266 4420823, San Luis, Argentina*

*{npasinetti, grodriguez, ymalvarado, jmfer, rag}@unsl.edu.ar*

## Abstract

Es usual que el aprendizaje de las nuevas tecnologías por parte de usuarios inexpertos presente desafíos debido a factores humanos y ergonómicos; es por ello que actualmente existe una tendencia a facilitar la interacción de los usuarios con dichas tecnologías a través de medios más intuitivos tales como comandos de voz, pantallas táctiles, sensores de movimiento, etc.

La interacción humano-computadora es fundamentalmente una tarea que implica procesar información. Para que la interacción entre el usuario y la computadora sea lo suficientemente eficiente, la interfaz debería ser diseñada de acuerdo a las capacidades de procesamiento de información del usuario. En consecuencia, uno de los retos más importantes radica en preservar las vías naturales de comunicación, interacción y entendimiento de las personas proveyendo a las computadoras de una visualización e interacción lo más humana posible. Adicionalmente, se debe tener en cuenta que los medios de interacción deben ser usables, disponer de una apariencia gráfica sencilla y responder al usuario en tiempo real.

Esta propuesta de trabajo pretende combinar diversos medios de interacción natural con el fin de lo-

grar una comunicación avanzada entre un humano y una representación por computadora que denominaremos personaje virtual.

**Palabras Claves:** Realidad Virtual, Computación Gráfica, Interfaces Humano-Computadoras, Vida Artificial, Personajes Virtuales.

## Contexto

La propuesta de trabajo se lleva a cabo dentro de la línea de Investigación “Procesamiento de Información Multimedia” del proyecto “Tecnologías Avanzadas Aplicadas al Procesamiento de Datos Masivos”. Este proyecto es desarrollado en el ámbito del Laboratorio de Computación Gráfica de la Universidad Nacional de San Luis.

## 1 Introducción

Como una de las nuevas tecnologías emergentes, la Realidad Virtual (RV) permite la generación de entornos de interacción que facilitan nuevos contextos de intercambio y comunicación de información. Una aplicación de Realidad Virtual consiste en la inmersión del usuario en un ambiente generado por computadora mediante el cual se simula la realidad utilizando dispositivos interactivos que permiten enviar y recibir información [1, 2]. Usualmente, esta clase de aplicaciones ofrecen una experiencia en la cual el usuario se retroalimenta con

estímulos sintéticos que se corresponden con uno o más de uno de sus sentidos [3].

Las aplicaciones de Realidad Virtual a menudo realizan la interacción mediante los dispositivos clásicos tales como joystick, teclado, mouse, entre otros; de la misma manera que lo hacen los videojuegos. Estos dispositivos de interacción funcionan bien para personas expertas en informática y/o jugadores de videojuegos (“*gamers*”), sin embargo son un gran obstáculo en la interacción para muchos de los usuarios finales, los cuales a menudo no están familiarizados con este tipo de dispositivos [4, 5].

Esto ha dado lugar a la necesidad de evolución de los dispositivos existentes, obligando el surgimiento de nuevos dispositivos los cuales deben ser simples e intuitivos, de manera que permitan mejorar la interacción de los usuarios con los entornos de Realidad Virtual. Es en este contexto que surgen las *interfaces de usuario naturales* como una solución atractiva a la necesidad planteada [6]. No obstante, si bien las posibilidades que ofrecen las nuevas tecnologías son mucho mayores en comparación a aquellas que se ofrecían en décadas anteriores, los dispositivos de interacción tradicionales son aún altamente valorados y populares debido al soporte que ofrecen [7, 8].

La interacción inteligente humano-computadora es un campo de investigación en desarrollo que tiene por objetivo proveer al hombre de vías naturales de comunicación con una computadora, a modo de convertirlas en herramientas de ayuda. El supuesto colectivo establece que para que una computadora sea capaz de interactuar con un humano ésta debe tener las capacidades de comunicación de los humanos [9, 10]. Para ello es necesario realizar un análisis exhaustivo de las habilidades y formas de comunicación del ser humano, las cuales pueden resumirse en dos grandes grupos: comportamientos *verbales* y *no verbales*. Las palabras que pronunciamos y la entonación del discurso representan la parte verbal de la comunicación, mientras que la parte no verbal está constituida por un amplio conjunto de comportamientos tales como los gestos de las

manos, la expresión facial, la postura corporal, la orientación y la mirada, entre otros [11, 12].

Como consecuencia, las interfaces naturales intentan dotar al usuario de una interacción inteligente consistente en el uso de acciones intuitivas tales como la incorporación de canales *gestuales, posicionales, sonoros y biométricos*; permitiendo que las interfaces de usuario naturales sean invisibles al usuario debido a que reducen la carga cognitiva requerida para realizar la interacción. Es importante que las interfaces de usuario naturales resulten fáciles de aprender y que el proceso de aprendizaje sea rápido tanto para principiantes como para expertos. Además, resulta evidente que la interfaz debe brindar soportes a las tareas específicas requeridas por el usuario. En resumen las interfaces de usuario naturales deberían ser fuertemente expresivas, transparentes, flexibles y eficientes; proporcionando un marco para un lenguaje y diálogo multimodal, en vez de un marco para el simple intercambio de información [13, 14, 15, 16].

Paralelamente, los *personajes virtuales* son un tipo emergente de interfaz humano-computadora, que poseen una personificación y habilidades de interacción multimodal. Exhiben un aspecto similar al del ser humano, tanto en apariencia como comportamiento, son capaces de expresar estados emocionales, rasgos de personalidad y diferentes habilidades de conversación, pudiendo así reconocer y responder a canales verbales y no verbales. Como interfaz prometen incrementar la calidad de la comunicación entre los seres humanos y las computadoras, ya que están diseñados para comunicarse e interactuar de una forma similar a la humana [17, 18]. Se debe tener en cuenta que la calidad de una conversación ‘natural’ depende del número de canales o modalidades utilizados para intercambiar información con el usuario. Por lo tanto, para lograr un comportamiento realmente vivo se requiere que múltiples canales del personaje estén reproduciéndose simultáneamente [19, 20].

Adicionalmente, los personajes virtuales requieren de un ambiente que les otorgue un mayor grado de realismo y credibilidad. Por

ello es posible enmarcarlos en entornos inmersivos como los tipo CAVE (*Cave Automatic Virtual Environment*), los cuales ofrecen una visualización y apariencia, en tiempo real, más familiarizada con la realidad del ser humano; a la vez, este tipo de ambientes fomenta la implementación de múltiples canales de interacción avanzados y naturales al humano.

## 2 Líneas de Investigación y Desarrollo

En función de lo anteriormente expresado, un personaje virtual como interfaz de interacción debería incluir el uso de diferentes canales de comunicación, verbales y no verbales, capaces de provocar una sensación de entidad viva. De acuerdo al enfoque y complejidad del comportamiento implementado en un personaje, se pueden definir las siguientes líneas de investigación [21]:

- *Modalidad Sonora.* Por un lado, implica la simulación por computadora de la capacidad inherente en los humanos de escuchar y hablar, en términos computacionales significa capturar y comprender la voz hablada del usuario. El proceso de comprensión involucra procesar el significado de lo hablado y generar una respuesta interna, lo cual, dependiendo del grado de coherencia y complejidad deseado puede incluir desde complejos algoritmos de Inteligencia Artificial hasta Búsqueda Semántica. Posteriormente la respuesta interna debe ser expresada mediante el lenguaje hablado realizando la correspondiente síntesis de voz por computadora [22]. Por otro lado, puede incluir sonidos ambientales correspondientes al entorno virtual en el que se encuentre inmerso el personaje tales como el ruido de tránsito, el sonido de un bosque, entre otros [23].
- *Modalidad Gestual.* Dentro de la comunicación no verbal, los gestos realizados por distintas partes del cuerpo hu-

mano así como también del cuerpo completo son una herramienta natural que forma parte del ser humano. El desarrollo de interfaces basadas en gestos es otro medio de comunicación humano-computadora. Dichas interfaces podrían solucionar y dar soporte a problemas de distintas índoles, que van desde la interpretación de lengua de señas, control de *drones* mediante un protocolo gesticular, hasta la asistencia en entrenamientos físicos y atención médica [24, 25, 26].

- *Modalidad Posicional.* En un sistema inmersivo de interacción humano-computadora, es de importancia conocer la posición en tiempo real de las personas dentro del ambiente en el que se encuentran, ya que su ubicación puede brindar información para diversos objetivos tales como la detección de caídas, el reconocimiento de rostro, el conteo de personas, la detección de personas, entre otros [27]. En particular un personaje virtual, necesita realizar un *'tracking'* en tiempo real del usuario con el cual interactúa, por ejemplo de la posición de la cabeza del usuario con el fin de lograr una comunicación cara a cara [28].

Si bien los estudios se centrarán en los personajes virtuales con modalidades en general, la tecnología y modelos a desarrollar podrían adaptarse perfectamente a aplicaciones específicas tales como el reconocimiento biométrico por medio de la voz, la geometría de la mano, las dimensiones corporales, entre otros.

## 3 Resultados obtenidos / esperados

El grupo de trabajo, además de pertenecer a un proyecto de investigación de la Universidad Nacional de San Luis, ha desarrollado tareas dentro del marco de un Proyecto ALFA III de la Comunidad Europea, denominado

GAVIOTA (Grupos Académicos para la Visualización Orientada por Tecnologías Apropriadas), en el que participa la UNSL en conjunto con otras universidades de America Latina y Europa.

En función de ello, se han desarrollado investigaciones para conocer el estado del arte a nivel mundial y principales enfoques, métodos y técnicas existentes en relación con el comportamiento autónomo de las entidades de un escenario virtual con el propósito de incorporar a futuro nuevos conceptos a los ya existentes, tanto matemáticos como basados en la física.

Como consecuencia, se ha desarrollado un personaje virtual del tipo pregunta-respuesta, el cual han sido contextualizado dentro de un sistema de RV, en particular, para la prevención de accidentes de tránsito urbano con escenarios experimentales de la ciudad de Concepción (Chile). Este trabajo se ha abordado en forma conjunta con la Universidad de Bio Bio (Chile).

Actualmente las acciones se encuentran focalizadas en la incorporación de nuevas habilidades al personaje a fin de avanzar en la comunicación entre el usuario y la representación por computadora.

## 4 Formación de Recursos Humanos

Los trabajos preliminares de estudio del arte han permitido la realización de trabajos de fin de carrera de la Licenciatura en Ciencias de la Computación, así como también la definición de un trabajo de tesis de Maestría en Ciencias de la Computación, actualmente en ejecución. Asimismo se ha obtenido una beca de finalización de carrera otorgada por la Secretaría de Ciencia y Técnica de la Fac. de Cs. Físico Matemáticas y Naturales de la UNSL.

## References

- [1] William R. Sherman and Alan B. Craig. *Understanding Virtual Reality: Interface, Application, and Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.
- [2] G.C. Burdea and P. Coiffet. *Virtual Reality Technology*. Number v. 1 in Virtual Reality Technology. Wiley, 2003.
- [3] A.B. Craig, W.R. Sherman, and J.D. Will. *Developing Virtual Reality Applications: Foundations of Effective Design*. Foundations of Effective Design Series. Elsevier Science, 2009.
- [4] P. Issacs, J. Shrag, and P. S. Strauss. The design and implementation of direct manipulation in 3d. ACM SIGGRAPH. 2002.
- [5] S. Stannus, D. Rolf, A. Lucieer, and W. Chinthammit. Gestural navigation in google earth. In *OZCHI*, pages 269–272, 2011.
- [6] J. Blake. *Natural User Interfaces in . Net*. Manning Pubs Co Series. Manning Publications Company, 2012.
- [7] O. Portillo-Rodriguez, C.A. Avizzano, A. Chavez-Aguilar, M. Raspolli, S. Marcheschi, and M. Bergamasco. Haptic desktop: The virtual assistant designer. In *Mechatronic and Embedded Systems and Applications, Proceedings of the 2nd IEEE/ASME International Conference on*, pages 1–6, Aug 2006.
- [8] W. Hinoshita, T. Ogata, H. Kozima, H. Kanda, T. Takahashi, and H.G. Okuno. Emergence of evolutionary interaction with voice and motion between two robots using rnn. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 4186–4192, Oct 2009.
- [9] Michael Lew, ErwinM. Bakker, Nicu Sebe, and ThomasS. Huang. Human-computer intelligent interaction: A survey. In Michael Lew, Nicu Sebe, ThomasS. Huang, and ErwinM. Bakker, editors, *Human-Computer Interaction*, volume 4796 of *Lecture Notes in Computer Science*, pages 1–5. Springer Berlin Heidelberg, 2007.
- [10] J.A. Jacko. *Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction: 13th International Conference, HCI International 2009, San Diego, CA, USA, July*

- 19-24, 2009, *Proceedings*. Human-computer Interaction: 13th International Conference : Proceedings. Springer, 2009.
- [11] S. Kopp and I Wachsmuth. Model-based animation of co-verbal gesture. In *Computer Animation, 2002. Proceedings of*, pages 252–257, 2002.
- [12] A. Kendon. *Gesture: Visible Action as Utterance*. Gesture: Visible Action as Utterance. Cambridge University Press, 2004.
- [13] M. Roupé, P. Bosch-Sijtsema, and M. Johansson. Interactive navigation interface for virtual reality using the human body. *Computers, Environment and Urban Systems*, 43(0):42 – 50, 2014.
- [14] S. Kean, J. Hall, and P. Perry. *Meet the Kinect: An Introduction to Programming Natural User Interfaces*. Apress Series. Apress, 2011.
- [15] Christian Benoit, Jean-Claude Martin, Catherine Pelachaud, Lambert Schomaker, and Bernhard Suhm. Audio-visual and multimodal speech systems. *Handbook of Standards and Resources for Spoken Language Systems-Supplement*, 500, 2000.
- [16] D. Wigdor and D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Elsevier Science, 2011.
- [17] M. Mancini and C. Pelachaud. Generating distinctive behavior for embodied conversational agents. *Journal on Multimodal User Interfaces*, 3(4):249–261, 2009.
- [18] J. Cassell. *Embodied Conversational Agents*. MIT Press, 2000.
- [19] Nicu Sebe, MichaelS. Lew, and ThomasS. Huang. The state-of-the-art in human-computer interaction. In Nicu Sebe, Michael Lew, and ThomasS. Huang, editors, *Computer Vision in Human-Computer Interaction*, volume 3058 of *Lecture Notes in Computer Science*, pages 1–6. Springer Berlin Heidelberg, 2004.
- [20] D. Perez-Marin and I. Pascual-Nieto. *Conversational Agents and Natural Language Interaction: Techniques and Effective Practices*. Premier reference source. Information Science Reference, 2011.
- [21] Andrew D. Wilson. Sensor- and recognition-based input for interaction. In *The Human Computer Interaction Handbook*, chapter 7, pages 133–156. 2012.
- [22] Jennifer Lai, Clare-Marie Karat, and Nicole Yankelovich. Conversational speech interfaces and technologies. In *The Human Computer Interaction Handbook*, chapter 19, pages 381–392. 2008.
- [23] D.R. Begault. *3-D Sound for Virtual Reality and Multimedia*. AP Professional, 1994.
- [24] Orlando Erazo and Roberto Pico. Interfaces de usuario basadas en gestos manuales sin contacto para la sala de clases: una revisión bibliográfica. *Enfoque*, 2014.
- [25] Hiroo Iwata. The human-computer interaction handbook. chapter Haptic Interfaces, pages 206–219. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 2003.
- [26] C. Camporesi and M. Kallmann. A framework for immersive vr and full-body avatar interaction. In *Virtual Reality (VR), 2013 IEEE*, pages 79–80, March 2013.
- [27] Anh Tuan Nghiem, E. Auvinet, and J. Meunier. Head detection using kinect camera and its application to fall detection. In *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*, pages 164–169, July 2012.
- [28] M. L. Yuan, Gim Guan Chua, Farzam Farbiz, and Susanto Rahardja. Eye contact with a virtual character using a vision-based head tracker. In Erol Gelenbe, Ricardo Lent, Georgia Sakellari, Ahmet Sacan, Ismail Hakki Toroslu, and Adnan Yazici, editors, *ISCIS*, volume 62 of *Lecture Notes in Electrical Engineering*, pages 217–225. Springer, 2010.