



Revista Facultad de Ingeniería

ISSN: 0717-1072

facing@uta.cl

Universidad de Tarapacá

Chile

Herrero, Pilar; Antonio, Angélica de  
Diseño de un modelo de percepción para Agentes Virtuales Inteligentes basado en el sistema de  
percepción de los seres humanos  
Revista Facultad de Ingeniería, vol. 11, núm. 1, 2003, pp. 11-24  
Universidad de Tarapacá  
Arica, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=11411202>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

## DISEÑO DE UN MODELO DE PERCEPCIÓN PARA AGENTES VIRTUALES INTELIGENTES BASADO EN EL SISTEMA DE PERCEPCIÓN DE LOS SERES HUMANOS

Pilar Herrero<sup>1</sup> Angélica de Antonio<sup>1</sup>

Recibido el 08 de julio de 2003, aceptado 22 de septiembre de 2003

### RESUMEN

Una prioridad en Sistemas de Agentes Virtuales Inteligentes (AVIs) y Sistemas Multi-Agentes Virtuales Inteligentes (mAVIs), es introducir una percepción sensitiva similar a la que poseen los seres humanos que permita que los AVIs perciban, en tiempo real, el entorno que le rodea con una claridad de percepción similar a la que poseen los seres humanos. Algunos estudios dentro del área de investigación cognitiva introducen la percepción humana como el primer nivel de un modelo de “awareness” (consciencia). Basándonos en estos estudios, se ha desarrollado un modelo de percepción basado en uno de los más conocidos modelos de “awareness” desarrollados para trabajo colaborativo (CSCW). Este modelo es conocido como “Modelo Espacial de Interacción” (MEI) y gobierna “awareness” en Entornos Virtuales Colaborativos (EVC).

El modelo de percepción desarrollado extiende los conceptos fundamentales del MEI, introduciendo algunos factores propios de la percepción tanto visual como auditiva en humanos (como son la Agudeza Sensitiva o la Región de Transición Sensorial) y reinterpreta estos conceptos clave con el fin de introducirlos en el sistema perceptivo de un AVI. Una vez formalizado el modelo, se ha seleccionado una arquitectura de agentes para poder introducir, usar y validar este modelo. De entre todas las posibles arquitecturas, se ha seleccionado el Modelo de Triple Torre de Nilsson. La introducción de este modelo de percepción dentro de la arquitectura de agente seleccionada ha requerido la creación de un “motor de percepción” y un diseño semi-centralizado de percepción que permita manejar no sólo la percepción sensitiva del agente sino también algunos factores externos y adicionales tales como la atenuación.

Palabras claves: Entornos virtuales, agentes virtuales inteligentes, percepción, consciencia, Focus, Nimbus

### ABSTRACT

*A priority in Intelligent Virtual Agents Systems (IVAS) and multi- Intelligent Virtual Agents Systems (mIVAS) is to introduce human-like sensitive perception which allows an IVA to perceive its environment and surrounding objects in real-time with a human-like clarity of perception, gaining knowledge about its surroundings and experiencing any virtual simulated environment. Within the cognitive research area, there are some studies underwriting that human perception can be understood as a first level of an “awareness model”. Following on from this research, we have developed a human-like perceptual model based on one of the most successful awareness models in Computer Supported Cooperative Work (CSCW), called the Spatial Model of Interaction (SMI), which manages awareness in Collaborative Virtual Environments (CVEs) through a set of key concepts. The perceptual model developed, extends the key concepts of the SMI by introducing some human-like factors typical of visual and hearing human perception - like Sense Acuity or Sense Transition Region - which are reinterpreted with the aim of introducing them as the key concepts of an IVA human-like perceptual model.*

*Having formalised the perceptual model, an appropriate agent architecture on which to use and test this perceptual model has been selected: a vertical layered architecture known as the Nilsson’s Triple Tower Model. The introduction of this perceptual model into the agent perceptual block, has required the creation of a perceptual engine and the design of a semi-centralized approach to manage not just the sensitive agent perception, but also some other external and additional factors, such as attenuation.*

*Keywords: Virtual environments, intelligent virtual agents, perception, awareness, Focus, Nimbus*

<sup>1</sup> Universidad Politécnica de Madrid, Facultad de Informática Campus de Montegancedo s/n, 28.660 Boadilla del Monte – Madrid, [pherrero@fi.upm.es](mailto:pherrero@fi.upm.es), [angelica@fi.upm.es](mailto:angelica@fi.upm.es)

## INTRODUCCIÓN

Los entornos o mundos virtuales interactivos proporcionan un poderoso medio para el aprendizaje y el entrenamiento experimental, siendo casi ilimitado el conjunto de mundos que la gente puede explorar, periodos de tiempo que el usuario puede atravesar, situaciones que el usuario puede vivir, así como grados de interactividad que el usuario puede experimentar.

De hecho, lo que en realidad hace un entorno virtual (EV) distinto de todas las interfaces hombre-máquina convencionales es que el usuario siente la ilusión de estar totalmente rodeado de información, adquiriendo consciencia de otros usuarios que podrían estar conectados y teniendo un sentimiento de inmersión, presencia y un grado de interactividad (con otros participantes) muy elevado.

Como una consecuencia de este elevado sentimiento de inmersión cuando un usuario está navegando a través de un entorno virtual, consciente o inconscientemente, está esperando tener la posibilidad de vivir allí cualquier tipo de situaciones. Es más, el usuario no está preparado para distinguir qué clase de situaciones pueden ser vividas en un entorno virtual de aquellas que sólo podría vivir en la vida real.

En este intento de simular las apariencias y la funcionalidad de la vida real, es inevitable llegar al punto de simular al propio ser humano. A menudo, los entornos virtuales incorporan agentes inteligentes con forma humana y varios grados de inteligencia, obteniendo lo que se llama *Agentes Virtuales Inteligentes* (AVI).

El campo de los entornos virtuales inteligentes es relativamente nuevo, aunque crece rápidamente, y se caracteriza por su variedad. Es precisamente esta característica la que consigue que muchas otras áreas de investigación, como por ejemplo psicología o medicina, se sientan especialmente atraídas por este tipo de entornos, queriendo contribuir a su desarrollo.

Otro tipo de sistemas, denominados Sistemas Multi-Agente (SMA), contienen múltiples agentes que pueden actuar cooperativamente dependiendo de sus propios intereses.

El alcance de esta investigación abarca lo que se ha denominado sistemas mAVI-EV (EV habitados por múltiples-AVIs). Un mAVI-EV es un tipo particular de SMA en el cual tanto usuarios como agentes están representados físicamente dentro del entorno, pudiendo interactuar los unos con los otros.

Los sistemas mAVI-EV pueden ser utilizados para simular situaciones propias de la vida real desde cualquier punto de vista. Uno de los ejemplos más representativos son las simulaciones de guerra, donde situaciones de riesgo y entrenamientos peligrosos pueden ser simulados con este tipo de aplicaciones sin necesidad de tener que exponer a ningún ser humano a ninguna situación peligrosa. En este tipo de entrenamientos el realismo es muy importante, ya que cuanto más realistas sean las situaciones que se le plantean a los soldados, mejor preparados estarán para superarlas en la vida real.

Muchos trabajos en sistemas mAVI-EVs tienen como objetivo proveer a los agentes con un alto grado de realismo en cuanto a su representación física o incluso su comportamiento. Pero, sin embargo, muy pocos estudios se han realizado centrándose en la percepción.

El resultado de un AVI percibiendo un objeto/agente es llegar a tener consciencia de dicho objeto/agente. Este trabajo pretende dotar a los AVIs con mecanismos de percepción que les permitan tener un conocimiento “realista” de su entorno. Para ello proponemos un modelo de percepción para AVIs que introduzca una mayor coherencia con el sistema de percepción humano, incrementando de forma indirecta la coherencia psicológica del EV con la vida real.

Esta coherencia es especialmente importante si se quieren simular situaciones reales como, por ejemplo, programas de entrenamiento militar, donde los soldados deben ser entrenados para vivir y sobrevivir a todo tipo de experiencias, incluyendo situaciones de alto riesgo. Así pues, para obtener un entrenamiento útil con agentes, cada uno de éstos debería poseer un modelo de percepción similar al que tendría uno de estos militares en la vida real para que pudieran reaccionar a los mismos estímulos como si fueran soldados humanos. De lo contrario podrían ocurrir situaciones surrealistas, como por ejemplo que escucharan o vieran cosas que están demasiado lejos para ser percibidas o que se encuentran ocultas detrás de un obstáculo u otro objeto. El modelo de percepción que propone este trabajo introduce estas limitaciones dentro del modelo de percepción del agente con el propósito de modelar un sistema perceptivo para el agente que sea similar al que poseen los seres humanos en la vida real.

Después de haber indagado sobre cómo podríamos llevar a cabo esta idea para que un AVI tuviera “consciencia” de su entorno, encontramos que el concepto de “consciencia” es muy familiar dentro del campo del Trabajo Colaborativo (Computer Supported Co-operative Work, también conocido por las siglas CSCW).

Probablemente la mejor definición de “consciencia” en CSCW fue proporcionada por [7] en su seminario de coordinación en espacios de trabajo compartidos. Ellos definieron el concepto de “consciencia” como “una comprensión de las actividades de los demás, las cuales proporcionan un contexto para tu propia actividad”. Pero éste es un tema de investigación que hoy todavía está abierto a nuevas aportaciones y puntos de vista.

Diferentes entornos físicos han sido desarrollados para soportar la naturaleza pública del trabajo y promover su coordinación. Así pues, cuando el entorno es un entorno virtual en tres dimensiones de carácter semi-distribuido, dentro del cual el usuario siente un alto grado de inmersión, éste se conoce normalmente como Entorno Virtual Colaborativo (EVC). Una buena definición para EVC es la dada por Greenhalgh como “entorno virtual multi-usuario distribuido que permite trabajar y jugar de forma cooperativa”.

De hecho, en este tipo de entornos existen, desde hace tiempo, modelos de consciencia que han sido testados con resultados bastante satisfactorios [11]. Quizás el más conocido de todos ellos es el “Modelo Espacial de Interacción” (MSI) [1], [2], desarrollado para sistemas multi-usuario de CSCW, y éste es el que se decidió tomar como fuente de inspiración.

Siguiendo los estudios realizados por Endsely [8], [9], sobre “consciencia”, la percepción humana puede ser entendida como un primer nivel de un modelo de awareness. Así pues, en este trabajo se ha tomado el Modelo Espacial de Interacción (MSI) como punto de partida a la hora de diseñar este modelo de percepción.

El Modelo Espacial de Interacción fue el primer modelo de “consciencia” propuesto, en 1993, como resultado de las investigaciones dirigidas por el profesor S. Benford en la Universidad de Nottingham, y de las investigaciones dirigidas por L. Fahlén en el Swedish Institute of Computer Science. El modelo espacial se diseñó de forma que permitiera obtener información del entorno que rodea al usuario en entornos colaborativos, a través de una serie de mecanismos como son el focus, nimbus o la consciencia, que fueron definidos para ese modelo y que permiten gobernar la interacción en un entorno virtual.

La “consciencia” cuantifica el grado, la naturaleza o la calidad de la interacción entre dos objetos. Esta relación de consciencia entre *cada par* de objetos es *unidireccional*. Existen diferentes tipos de consciencia: *Consciencia Total*, *Consciencia Parcial* y *Sin Consciencia*. El “Focus” determina la zona del espacio desde la cual un objeto recibe información, y el “Nimbus” determina la zona del espacio en la cual un objeto proyecta su información. Tanto el focus como el

nimbus como la consciencia son *específicos* para cada medio de transmisión de información.

Aunque las primeras implementaciones experimentales del modelo se realizaron en 1994, se podría decir que el primer prototipo experimental que realmente incluye, casi en su totalidad, el modelo espacial de interacción es MASSIVE. Este modelo fue inicialmente demostrado en el sistema DIVE y posteriormente implementado en su totalidad en el sistema MASSIVE. MASSIVE ha sido desde entonces utilizado para llevar a cabo numerosos encuentros virtuales, y de su estudio y desarrollo han surgido una gran cantidad de publicaciones.

El Modelo Espacial de Interacción ha sufrido varias extensiones desde sus orígenes. Un ejemplo son las propuestas por [24]. Pero quizá la más conocida de todas ellas es la propuesta por Greenhalgh en 1997 y conocida como “Third Party Objects” [10].

Sin embargo en este trabajo se va a partir del Modelo Espacial de Interacción en su forma básica original, ya que es el modelo más abstracto, abierto y sencillo de remodelar de acuerdo con los propósitos de este trabajo de investigación. Partiendo de los conceptos básicos definidos en este modelo vamos a diseñar un modelo de percepción para AVIs que permita dotarles con una claridad de percepción similar a la que tenemos los seres vivos en la vida real.

Así pues, el primer objetivo de esta investigación será realizar una reinterpretación de este modelo (desarrollado para CSCW) que permita aplicarlo a AVIs, que sirva como modelo de percepción para AVIs, y que introduzca una serie de factores propios de la percepción humana en la vida real.

## PERCEPCIÓN EN AGENTES INTELIGENTES

La combinación de mundos virtuales interactivos y agentes inteligentes ha experimentado un excitante progreso, permitiendo generar nuevas herramientas con diferentes fines y aplicaciones. Hoy en día, el usuario puede participar incluso interactivamente en una escena virtual, teniendo la posibilidad de experimentar con criaturas autónomas en entornos virtuales simulados sobre potentes máquinas.

Los AVIs reciben información del entorno que les rodea por medio de sensores virtuales que pueden ser visuales, táctiles o auditivos [22], [29]. Un AVI puede simplemente evolucionar dentro del entorno, puede interactuar con el propio entorno o incluso puede comunicarse con otros AVIs que se encuentren dentro del mismo entorno. Así pues en este tipo de entornos la

percepción del entorno que rodea a los agentes juega un papel muy importante.

La idea de percepción sensorial en agentes entendida como un proceso de búsqueda de información dentro del entorno tiene sus orígenes en la psicología humana. En 1994 Tu y Terzopoulos [31] introdujeron un innovador sistema de percepción artificial para peces, sistema que posteriormente fue perfeccionado por el propio Terzopoulos junto con Rabie [25], [26], [27], [28] hasta proponer un novedoso paradigma de percepción visual activa para animales, donde éstos, autónomamente, controlaban sus ojos y sus músculos, y mediante algoritmos de visión computacional analizaban continuamente las imágenes que llegan a la retina.

Posteriormente, Blumberg [3], introdujo sensores sintéticos dentro de un perro animado y autónomo, generando un sistema de percepción artificial llamado ALIVE, que podía interactuar con su propio entorno.

Aunque hace unos años el propósito de los sistemas de percepción diseñados para agentes era básicamente recopilar información básica sobre el entorno que rodeaba al propio agente, actualmente estos requisitos han cambiado, orientándose hacia otro tipo de aplicaciones más realistas que requieren información fiable y detallada del entorno que rodea al agente.

Este tipo de aplicaciones tienen como finalidad la simulación y el entrenamiento de agentes especializados para vivir situaciones reales que pueden ser arriesgadas para el usuario.

El desarrollo de estos modelos realistas ha sido calificado de “importante” por CGFs (Computer Generated Forces), haciendo un gran énfasis en el desarrollo de percepción y comportamiento “humanos” [17]. Se han desarrollado varios modelos de percepción y acción “humanos”, como por ejemplo “*The Model Human Processor*” (MHP) [4], que probablemente es el modelo humano de procesamiento de la información más referenciado a nivel mundial, o “*The Executive Process-Interactive Control*” (EPIC), que ha sido utilizado para trabajar con información humana sensorial de bajo nivel.

Además, los AVIs deben estar programados con un conjunto de reglas de comportamiento que les permita reaccionar a su entorno, tomando decisiones basadas no sólo en la información que reciben del entorno, sino también en su propia memoria y razonamiento.

Los AVIs han sido desarrollados con varios propósitos y características y por este motivo la percepción en estos agentes ha sido modelada de varias formas. Dependiendo de la finalidad para la cual éstos agentes

hayan sido diseñados su modelo de percepción puede centrarse más en el procesamiento de las entradas sensoriales o en el proceso cognitivo de percepción.

El procesamiento de las entradas sensoriales del agente puede llevarse a cabo de varias formas. Algunos prefieren utilizar sensores virtuales, que pueden ser visuales, táctiles o auditivos [29], [30]. Un ejemplo de este tipo de percepción sensorial en agentes son los “pilotos virtuales” descritos en 1999 por Hill [16], [17], donde la apariencia física del agente o el control de su cuerpo no juegan un papel tan importante como su sistema de percepción.

Los pilotos virtuales desarrollan tareas tales como volar siguiendo una ruta y buscando vehículos enemigos. Estas tareas requieren comportamientos tales como búsqueda, fijación o seguimiento que requieren la coordinación de la parte perceptiva y cognitiva del piloto así como del motor del sistema. El funcionamiento de estos pilotos requiere una componente de “atención” propia de los seres humanos.

En cuanto a la implementación del proceso cognitivo de percepción, éste puede centrarse sólo en la implementación de un factor (p.e. atención), en la implementación de una parte del proceso cognitivo (p.e. reconocimiento o aprendizaje), o incluso en el modelado de la arquitectura cognitiva [6]. Algunos de los métodos empleados para la implementación del proceso cognitivo de percepción son:

- Técnicas de “representación” cognitiva, como, por ejemplo, las empleadas por Hill et al. [18], [20]. Esta implementación está basada en el “marco” propuesto por Yeap y Jefferies [32] y que representa un entorno local como una estructura llamada “Representación Espacial Absoluta”.
- Redes neuronales para reconocer y predecir patrones de comportamiento. Este método ha sido utilizado por Hill [19] para poder llevar a cabo alguna tarea específica en alguna de sus aplicaciones. Hill combina redes de neuronas con el modelo perceptual diseñado para los “pilotos virtuales” [33], [17], creando un sistema híbrido que además está dotado con la capacidad de razonamiento propia de un agente “Soar” [16].
- Modelos de “atención” visual para “humanos virtuales” [5], [17]. Chopra se basa en investigaciones sobre la psicología humana, especificando el tipo de atención requerida para una gran variedad de tareas básicas, como puede ser la manipulación de un objeto, así como mecanismos para “dividir” la atención entre varios objetos que se perciben en el entorno. En el piloto virtual, Hill se

centra en proporcionar un modelo perceptual de atención para “humanos virtuales” en campos de batalla simulados [17], desarrollando una simulación distribuida e interactiva de combates, llamada ModSAF.

- Otra forma distinta de implementar la percepción es utilizada por Rickel et al. en Steve. Steve es un agente pedagógico (Soar Training Expert for Virtual Environments) que facilita el proceso de aprendizaje en procesos educativos y formativos [23]. Steve contiene tres módulos que se comunican entre sí por paso de mensajes: percepción, cognición y motor de control. El módulo de percepción visualiza los mensajes e identifica los eventos que son importantes para Steve. El módulo cognitivo interpreta la entrada que recibe del módulo de percepción (anteriormente descrito), enviando mensajes al módulo de percepción cuando este lista para actualizar el estado del entorno virtual. El módulo cognitivo además le envía al motor de control mensajes con comandos que deben realizarse y que el propio motor de control debe descomponer en comandos de bajo nivel para poder procesarlos correctamente.

Como se ha podido ver por medio de este breve resumen, los modelos de percepción que se han desarrollado hasta el momento para agentes virtuales inteligentes utilizan diferentes métodos y técnicas. Sin embargo, ninguno de ellos se ha concentrado en determinar con qué claridad un agente puede determinar un objeto ni como determinar si esta claridad de percepción sería la misma que la que tendría un humano situado en la misma posición e intentando percibir el mismo objeto.

## DISEÑO DE LA ARQUITECTURA DEL AGENTE

Para poder llevar a cabo este objetivo, y así poder dotar a mAVI-EVs con un sistema de percepción similar al que poseemos los seres humanos, necesitamos primeramente partir de una arquitectura de agentes sobre la cual introducir ese sistema de percepción.

Así pues, el punto de partida ha sido una arquitectura vertical de capas como la triple torre de Nilsson [21]. Esta arquitectura posee tres bloques (Fig. 1) que representan la percepción del agente, el procesamiento central del agente y el conjunto de acciones que realiza el agente, pero este trabajo centra toda su atención en el bloque de percepción, que se encarga del modelado del sistema de percepción del agente. Todos estos estudios se realizarán tanto para la percepción visual [14] como para la percepción auditiva [15], dejando el resto de los sentidos para posibles líneas de investigación futuras.

El bloque de percepción estará compuesto de tres módulos: *Percepción Sensitiva*, *Atenuación* y *Filtrado Interno*. El módulo de percepción sensitiva se encargará de percibir el entorno haciendo uso de un motor de percepción externo. Las sensaciones perceptivas que resultan de este módulo podrán verse atenuadas por una serie de factores. Un ejemplo de estos factores son el tiempo de exposición o el cansancio físico. El módulo de filtrado se encargará de seleccionar de entre todos los objetos presentes en el foco de percepción aquellos en los cuales el usuario está interesado.

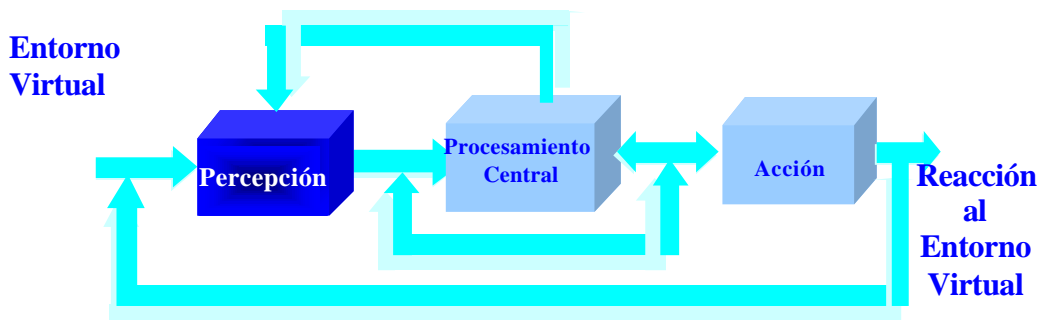


Fig. 1.- Arquitectura del agente

## FACTORES HUMANOS INTRODUCIDOS EN EL MODELO

Para poder hacer que este sistema de percepción sea más realista introduciremos algunos conceptos típicos del sistema de percepción humana, como por ejemplo [11], [12], [13], [14], [15]:

- *Agudeza Sensorial*: Medida de las habilidades de cada uno de los sentidos para resolver detalles finos. Este factor depende de la propia persona, del sistema de acomodación del propio sentido y del entorno que rodea a la persona mientras ésta percibe.
- *Zona de Transición Sensorial*: Aquella región del espacio dentro de la cual es posible percibir a alto nivel formas, sombras o sonidos, aunque no se pueda llegar a percibir detalles de dichos objetos/sonidos. La existencia de esta zona está basada en la gradualidad del sistema de percepción humana, existiendo siempre una zona de “transición gradual” entre la percepción total o máxima y la percepción nula.

Estos conceptos se introducirán de una manera “coherente” tomando como base el modelo de consciencia seleccionado y redefinido para AVIs con una única finalidad: modelar un sistema de percepción similar al que poseen los seres humanos para mAVI-EVs.

## FACTORES FÍSICOS INTRODUCIDOS EN EL MODELO

Además de los factores propios del sistema de percepción humano que han sido mencionados en el apartado anterior, en este modelo se han tenido en cuenta, de forma directa o indirecta, algunos factores físicos propios de la proyección de la imagen de un objeto o de la propagación de un sonido. Estos factores han sido: *el flujo de proyección* de energía irradiado por un objeto; la teoría del color; la directividad del sonido; el cono de confusión auditivo y las diferencias interaurales (en tiempo y en frecuencia) en la propagación del sonido.

## REINTERPRETACIÓN DEL MSI

Cuando el “Modelo Espacial de Interacción” MEI fue diseñado, éste no contempló ninguno de los factores físicos y humanos que se han introducido en esta memoria. Así pues, después de haber seleccionado los factores físicos y humanos a tener en cuenta en el modelo, se ha hecho una reinterpretación de los

conceptos principales de este modelo para que pueda incluir estos nuevos factores [14],[15].

## Percepción Visual

Se ha definido una nueva función matemática que representa el “focus” de percepción del agente para que éste sea similar al focus de percepción del ser humano (ecuación 1, Fig. 2) [14].

$$\begin{aligned} \mu_y \leq y \leq \mu_y + D_m \\ (x - \mu_x)^2 + (z - \mu_z)^2 &\leq (\tan(\theta))^2 * (y - \mu_y)^2 \\ (x - \mu_x)^2 + (z - \mu_z)^2 &\leq (\tan(\theta'))^2 * (y - \mu_y)^2 \end{aligned} \quad (1)$$

$$D_m \leq \frac{s}{\tan(\text{MAR})}$$

Donde:

- $(\mu_x, \mu_y, \mu_z) \rightarrow$  Representa la posición del ojo del agente en un sistema de referencia en 3D.
- $D_m$  : Representa la distancia a la cual el agente es capaz de resolver el objeto que está percibiendo.
- $(x, y, z)$  : Representa cualquier punto dentro del foco.
- Ángulos :
  - $\theta'$ : Representa el ángulo que delimita la región central de percepción.
  - $\theta$ : Representa el ángulo que delimita la región lateral de percepción.

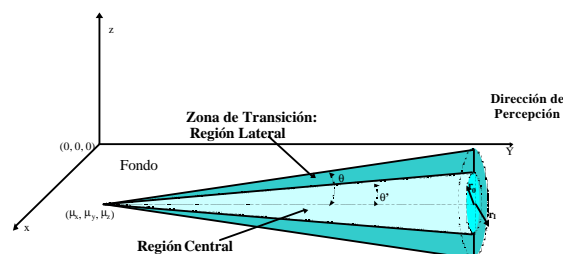


Fig. 2.- Focus físico con visión lateral (componente angular)

En la implementación del modelo se ha distinguido entre el focus general, que tiene una longitud infinita, y el focus asociado al agente, que tiene una longitud finita que depende las características del propio agente.

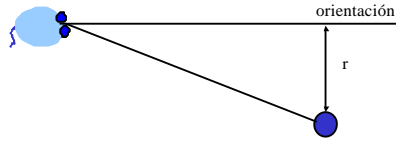


Fig. 3.- Orientación del ojo del agente y posición del objeto a percibir

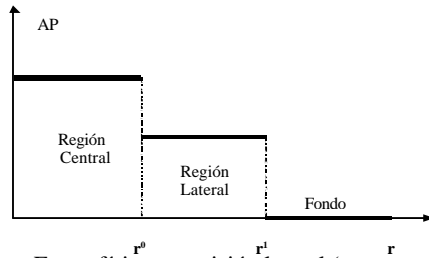


Fig. 4.- Focus físico con visión lateral (componente espacial)

Cuando un agente percibe un objeto en el entorno, la percepción es diferente dependiendo del área donde este objeto esté situado. Tanto la posición del objeto como la orientación del ojo del agente (Fig. 3) y la región donde está situado (Fig. 4) juegan un papel muy importante a la hora de determinar la claridad de percepción del objeto en el entorno. En la Fig. 4 el área de percepción (AP) indica si un objeto se encuentra dentro del focus y, si esto ocurre, dentro de que área se encuentra situado.

En la Fig. 2 se puede apreciar cómo se pueden distinguir dos conos de percepción: el *cono interno de percepción* (con ángulo  $\theta'$ ) y el *cono externo de percepción* (con ángulo  $\theta$ ). El cono interno de percepción representa el campo de visión central del agente sin tener en cuenta su visión lateral, mientras que el cono externo de percepción sí tiene en cuenta la visión lateral del agente.

Por otro lado, partiendo tanto de la definición original del nimbus de un objeto como de las implementaciones que Greenhalgh [10] ha realizado de este concepto en diferentes sistemas y aplicaciones, se ha representado el nimbus de un objeto por un elipsoide (Fig. 5, ecuación 2) o una esfera (ecuación 3), dependiendo de la cónica que circunscriba el objeto.

$$\left( \left( \frac{x - \mu_x}{a} \right)^2 + \left( \frac{y - \mu_y}{b} \right)^2 + \left( \frac{z - \mu_z}{c} \right)^2 \right) \leq 1 \quad (3)$$

$$(x - \mu_x)^2 + (y - \mu_y)^2 + (z - \mu_z)^2 \leq R^2 \quad (4)$$

Donde:

- $(\mu_x, \mu_y, \mu_z)$ : Representa el centro geométrico del objeto.
- $(a, b, c)$  : Representa los parámetros del elipsoide.
- $R$  : Representa el radio de la esfera (cuando  $a = b = c = R$ ).

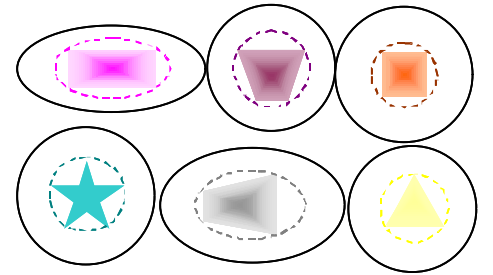


Fig. 5.- Representación física del Nimbus

### Percepción Auditiva

Cuando el ser humano recibe un sonido, éste puede llegar de dos formas distintas: combinando la información que le llega de los dos oídos "*binaural*" o tomando la información que le llega a cada uno de los oídos de forma independiente "*monaural*". Este trabajo ha centrado toda su atención en la percepción auditiva binaural.

La posición de una fuente de sonido relativa al centro de la cabeza del oyente se suele expresar en términos de dos ángulos: ángulo azimutal ( $\theta$ ), que se mueve en el rango  $[0^\circ, 360^\circ]$ , y ángulo de elevación ( $\phi$ ), que se mueve en el rango  $[-90^\circ, 90^\circ]$  (Fig. 6).

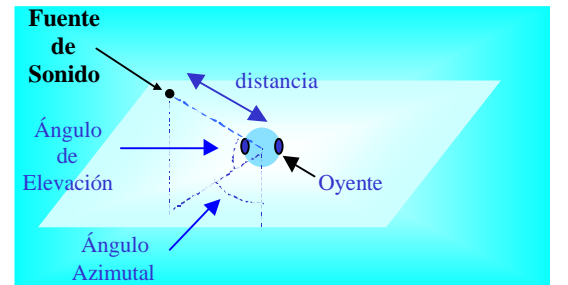


Fig. 6.- Posición relativa al oyente de una fuente de sonido

Los ángulos azimutal y de elevación varían independientemente, razón por la cual se ha preferido representar el focus auditivo de este modelo de



percepción por un doble cono elíptico (ec. 5, Fig. 7) [15].

$$\begin{aligned} \hat{i}_y - d \leq y \leq \hat{i}_y + d \\ \frac{(x - \hat{i}_x)^2}{b^2} + \frac{z^2}{a^2} \leq 1 \end{aligned} \quad (5)$$

$$a = d * \tan \varphi$$

$$b = d * \tan \theta$$

Donde:

- $(\mu_x, \mu_y, \mu_z)$  : Representa la posición del agente en un sistema de referencia en 3D.
- $d$  : Representa a la cual el agente puede percibir ese sonido. Esta distancia está asociada a los ángulos  $\theta$  y  $\varphi$
- $(x, y, z)$  : Representa cualquier punto dentro del foco.
- Ángulos :
  - $\theta$ : Representa el ángulo azimutal.
  - $\varphi$  Representa el ángulo de elevación.

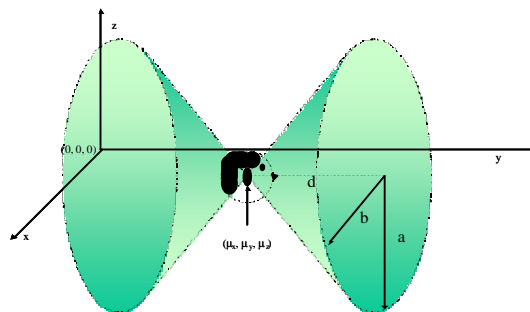


Fig. 7.- Focus auditivo (el agente se encuentra mirando hacia la derecha)

En cuanto al nimbus auditivo, de entre todos los factores que podrían afectar a la proyección del sonido en un medio, en este trabajo solamente se ha tenido en cuenta la directividad del sonido, dejando factores tales como la presencia de factores no-lineales o la homogeneidad del medio para futuras extensiones de este modelo.

Además este modelo se ha centrado en la proyección de la voz humana, para la cual existe una serie de patrones de directividad establecidos a diferentes frecuencias (Fig. 8).

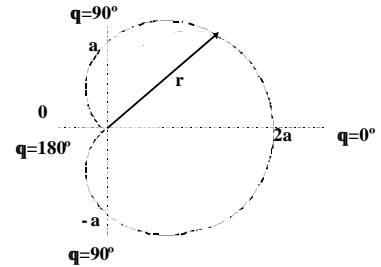


Fig. 8.- Nimbus asociado a la voz humana

La Fig. 8 representa el área dentro de la cual la voz humana se proyecta. Esta figura corresponde con un cardioide cuya ecuación en coordenadas polares es:

$$r = a(1 + \cos \theta) \quad (6)$$

Siendo la representación en coordenadas cartesianas de esta ecuación:

$$(x^2 + y^2 - ax)^2 = a^2(x^2 + y^2) \quad (7)$$

## CLARIDAD DE PERCEPCIÓN

El objetivo final de este modelo de percepción es poder determinar si un agente puede percibir un objeto que se encuentra en el entorno que le rodea. Para ello este diseño se ha concentrado en el bloque de percepción sensitiva introducido en la arquitectura del agente. Este bloque simula el proceso típico de percepción humana por el cual los receptores sensoriales (ojos y oídos en este caso) reciben sensaciones del entorno. La percepción sensitiva dependerá de algunos factores sensoriales relevantes. Estos factores pueden clasificarse en cuatro grupos:

- **Factores Humanos:** Dependen de la naturaleza del propio individuo. Entre este tipo de factores se encuentran la agudeza sensorial, la zona de transición sensorial y los filtros internos.
- **Factores Físicos:** Dependen de la disposición física del entorno. Entre estos de factores se encuentra la distancia entre el objeto, o la fuente de sonido, y la posición del agente.
- **Factores propios del objeto o de la fuente de sonido:** Dependen de la forma física del objeto y de sus propiedades. Entre estos de factores se encuentran el tamaño del objeto o la intensidad de la fuente de sonido.
- **Adaptadores:** Modifican la percepción que el agente puede tener del entorno.

La distancia entre el objeto, o la fuente de sonido, y la posición del agente, y la claridad de percepción, en general, son “conceptos clave” en el modelo de percepción, porque introducen más realismo, credibilidad y eficiencia. Así, sólo sería necesario conocer su valor para saber si un agente, a una determinada distancia, puede leer un cartel o escuchar un sonido.

La claridad de percepción es una medida de la habilidad que un agente posee para percibir un objeto/sonido que se encuentra presente en su *focus* de percepción. Para ello se ha utilizado el concepto de *consciencia* tal y como fue introducido en sus orígenes por el Modelo Espacial de Interacción. Así pues si la percepción que se realiza es visual, una vez que el *nimbus* del objeto que se desea percibir intersecciona con el *focus* del agente que percibe el entorno, la percepción sensitiva del agente entra en juego y calcula la claridad de percepción que el agente tiene de ese objeto/sonido. Pero si la percepción es auditiva el sonido debe propagarse, por medio de su *nimbus*, hasta alcanzar el *focus* del agente y además este sonido debe llegarle al propio agente con una intensidad que pueda ser percibida por el agente, es decir, la intensidad con la cual el agente recibe el sonido debe estar en el rango audible del agente.

Matemáticamente la claridad de percepción vendría dada por las funciones matemáticas:

- Medio visual (región central):

$$0.0 \leq d \leq d_1 \quad CP(d) = \lambda d$$

$$d_1 \leq d \leq d_2 \quad CP(d) = CP_{\max} \quad (8)$$

$$d \geq d_2 \quad CP(d) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \exp \left\{ -\frac{(d - d_2)^2}{2 \cdot \sigma^2} \right\}$$

siendo

- $d_1$  : La distancia mínima entre el objeto y el ojo para que dicho objeto pueda ser percibido.
- $d_2$  : La distancia máxima entre el objeto y el ojo para que dicho objeto pueda ser percibido con un elevado nivel de detalles.

- Medio visual (región lateral):

$$0.0 \leq d \leq d'_1 \quad CP(d) = \lambda d$$

$$d'_1 \leq d \leq d'_2 \quad CP(d) = CP_{\max} \quad (9)$$

$$d \geq d'_2 \quad CP(d) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \exp \left\{ -\frac{(d - d'_2)^2}{2 \cdot \sigma^2} \right\}$$

siendo:

$$d'_1 > d_1 \quad d'_2 < d_2$$

- Medio auditivo (percepción monaural):

$$CP_M(d) = CP_0(\theta) - 20 \cdot \log(d) \quad (10)$$

$$CP_0(\theta) = \begin{cases} 10 & \theta = 0^\circ \\ 20 & \theta = 30^\circ \\ 30 & \theta = 60^\circ \\ 40 & \theta = 90^\circ \end{cases} \quad (11)$$

siendo:

- $\theta$  : El ángulo azimutal (Fig. 6)
- $d$  : La distancia entre la fuente de sonido y el oído del agente
- Medio auditivo (percepción binaural):

$$CP_B(d) = \begin{cases} CP_M(d) \cdot \frac{r}{d} & d \in [0.15, 1] \\ -20 \cdot \log(d) & d \in [1, d_{\max}] \end{cases} \quad (12)$$

siendo:

- $\theta$  : El ángulo azimutal (figura 6)
- $d$  : La distancia entre la fuente de sonido y el oído del agente
- $r$  : El radio de la cabeza del agente que percibe dicho sonido

## IMPLEMENTACIÓN DEL MODELO

El lenguaje de programación elegido para llevar a cabo la implementación del modelo ha sido C++ y dicha implementación se ha centrado en la construcción de una librería independiente de cualquier sistema.

La validación de este modelo se ha desarrollado sobre una plataforma, la cual se ha llamado MASSIM\_AGENT (Fig. 9). Esta plataforma ha sido el resultado de la integración de la plataforma MASSIVE, desarrollada por la Universidad de Nottingham, y la herramienta para el desarrollo de agentes SIM\_AGENT, desarrollada por la Universidad de Birmingham.

Aunque se ha realizado la validación de este modelo sobre este sistema, el objetivo más inmediato es integrar este modelo junto con el sistema resultante de un proyecto llamado MAEVIF (Modelo para la Aplicación de Entornos Virtuales Inteligentes a la Educación y el Entrenamiento). Este proyecto está siendo financiado actualmente por el Ministerio de Ciencia y Tecnología (TIC00-1346).

La implementación del este modelo de percepción tiene otra importante característica que permite la optimización de los recursos del sistema. Aunque la mayoría de los sistemas mAVI tienen una implementación distribuida de su modelo perceptivo, en la cual cada agente es responsable de la percepción que tiene del entorno que le rodea, el sistema introduce una aproximación semi-centralizada del modelo de perceptivo del agente. Esto significa que cada agente es responsable solamente de una parte de su percepción mientras que el sistema es responsable del resto.

## VALIDACIÓN

Para poder validar el modelo de percepción que se ha diseñado e implementado, se ha introducido dicho modelo dentro de la arquitectura de agentes seleccionados (Fig. 1).

Después de haber dotado al agente con este modelo de percepción, se han calibrado los parámetros preceptuales del agente para que éstos sean similares a los que poseería un humano.

Posteriormente se ha planteado una serie de escenarios, en los cuales la percepción del agente debe ser lo más realista posible y se ha realizado una primera fase de evaluación basada en estos escenarios (Figs. 10 y 11). El propósito de esta validación era corroborar si los resultados obtenidos eran intuitivamente coherentes con aquellos que deberíamos obtener en la vida real.

A continuación se ha seleccionado un grupo de usuarios y se les ha sometido a una serie de experimentos con el fin de averiguar cuáles son sus preferencias y cuál es la credibilidad del modelo.

La valoración del grado de “claridad de percepción” ha sido realmente difícil para aquellos participantes que han tomado parte en el proceso, ya que este concepto es muy subjetivo y por lo tanto es muy difícil establecer cuál es el límite entre, por ejemplo, una claridad de percepción baja y muy baja.

Sin embargo, en conjunto, los resultados obtenidos han sido bastante satisfactorios y coherentes en la validación basada en escenarios, la desviación típica obtenida en los experimentos ha sido bastante baja (se han obtenido una desviación típica media cercana al 15%) y la reacción de los usuarios a este modelo de percepción ha sido satisfactoria a pesar de haber introducido modelos de razonamiento muy simples en estos agentes.

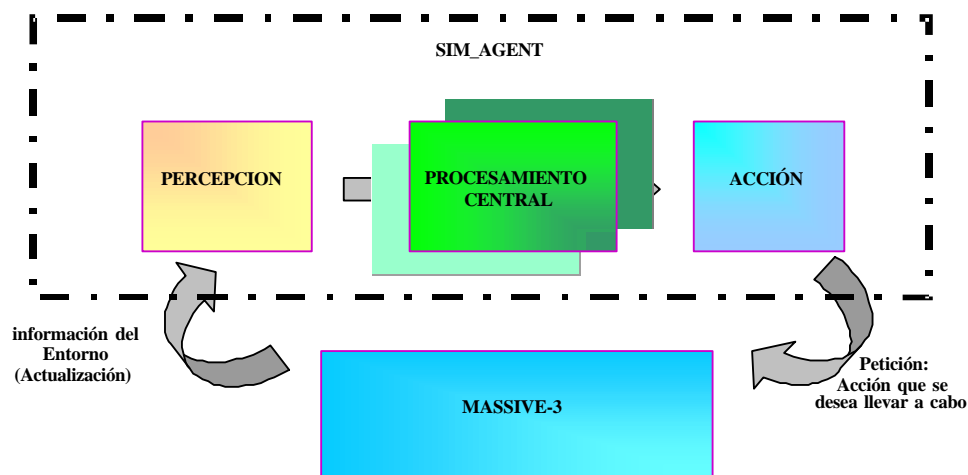


Fig. 9.- MASSIM\_AGENT: Comunicación



Fig. 10.- Escenario elegido para validar el modelo



Fig. 11.- Escenario elegido para validar el modelo

## CONCLUSIONES

En este trabajo se ha diseñado e implementado un modelo de percepción para Agentes Virtuales Inteligentes (AVI) basado en uno de los más conocidos modelos de “awareness” desarrollados para trabajo colaborativo (CSCW). Este modelo es conocido “Modelo Espacial de Interacción” (MEI) y gobierna “awareness” en Entornos Virtuales Colaborativos (EVC) [1]. De este modo, las capacidades perceptivas de un AVI serían similares a las de un humano.

Este modelo de percepción extiende los conceptos fundamentales del MEI, introduciendo algunos factores propios de la percepción en humanos, reinterpreta estos conceptos clave con el fin de introducirlos en el sistema perceptivo de un AVI e introduce las definiciones de algunos conceptos nuevos - como por ejemplo el flujo de proyección o la claridad de percepción - y las funciones matemáticas que modelan y describen tanto los conceptos fundamentales del MEI aplicados a AVI como los nuevos conceptos que aporta este modelo.

Este trabajo se ha centrado en los sentidos más utilizados en Entornos Virtuales (EV) por el momento - visión y oído - y en algunos factores característicos del sistema perceptivo humano como son la Agudeza Sensitiva o la Región de Transición Sensorial.

Este modelo de percepción ha sido introducido en una arquitectura de agentes genérica - como es el Modelo de Triple Torre de Nilsson - con el fin de poder validar y evaluar adecuadamente el modelo desarrollado, y para ello ha sido necesario la creación de un “motor de percepción” que permita manejar no sólo la percepción sensitiva del agente sino también algunos factores externos y adicionales tales como la atenuación. Este “motor de percepción” es centralizado y se complementa con otro; la propia “percepción del agente” que es distribuida.

El modelo de percepción resultante permite que un AVI perciba, en tiempo real, el entorno que le rodea con una claridad de percepción similar a la de un ser humano mediante un modelo semi-centralizado de percepción.

## AGRADECIMIENTOS

El trabajo presentado en este artículo ha sido avalado y parcialmente financiado por la Universidad de Nottingham, habiéndose desarrollado una gran parte de este trabajo allí - en el centro de investigación “Communication Research Group (CRG)” de la Escuela de Informática y Tecnología de la Información de dicha

universidad - bajo la dirección y continua supervisión de Steve Benford y Chris Greenhalgh.

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia y Tecnología Español a través del proyecto TIC00-1346

## REFERENCIAS

- [1] S. D. Benford and L. E. Fahlén; “A Spatial Model of Interaction in Large Virtual Environments”. Proceedings of the Third European Conference on Computer Supported Cooperative Work (ECSCW'93). Milano. Italy. Kluwer Academic Publishers, pp. 109-124, 1993.
- [2] S. Benford et al.; “User Embodiment in Collaborative Virtual Environments. Proceedings of ACM Conference on Human Factors in Computing Systems”, (CHI'95). Denver, Colorado, USA, pp. 242-249, 1995.  
[http://www.acm.org/sigchi/chi95/Electronic/documnts/papers/sdb\\_bdy.htm](http://www.acm.org/sigchi/chi95/Electronic/documnts/papers/sdb_bdy.htm)
- [3] B. Blumberg; Go with the Flow, “Synthetic Vision for Autonomous Animated Creatures. Proceedings of the First International Conference on Autonomous Agents”, (Agents'97). Marina del Rey. California, pp. 538-539, 1997.
- [4] S. Card, T. Moran, A. Newell; The psychology of human-computer interaction. Hillsdale. NJ. Lawrence Erlbaum Associates. Inc, 1983.
- [5] S. Chopra -Khullar and N. Badler; Where to look? Automating attending behaviors of virtual human characters. Autonomous Agents and Multi-agent Systems 4(1/2), pp. 9-23, 2001.
- [6] S. Das, D. Grecu COGENT; “Cognitive agent to amplify human perception and cognition, Proceedings of the Fourth International Conference on Autonomous Agents. Barcelona. Spain, pp. 443-450, 2000.
- [7] P. Dourish and V. Bellotti; “Awareness and Coordination in Shared Workspaces”,. Proceedings of the 4th ACM Conference on Computer-Supported Cooperative Work (CSCW). Toronto. Ontario. Canada, pp.107-114, 1992.
- [8] M. Endsley; “Design and evaluation for situation awareness enhancement”, Proceedings of the Human Factors Society and 32<sup>nd</sup> Annual Meeting, Santa Monica. CA: Human Factors Society, 1, pp. 97-101, 1988.

- [9] M. Endsley; "Towards a theory of situation awareness", Technical report, Texas Technical University, Department of Industrial Engineering, 1993.
- [10] C. Greenhalgh; "Large Scale Collaborative Virtual Environments", Doctoral Thesis. University of Nottingham, 1997.
- [11] P. Herrero; "Awareness of interaction and of other participants", Technical Report. Amusement Esprit Project 25197, 1999.
- [12] P. Herrero; A. De Antonio, J. Segovia; "Is the Awareness of Avatars in a Virtual World Different from Human Awareness?", Workshop on the Future of Cves: Voltage in the Milky Night: The Future of CVE's. The Third International Conference on Collaborative Virtual Environments. ACM Collaborative Virtual Environments 2000. San Francisco. California. USA, 2000.
- [13] P. Herrero, A. De Antonio, S. Benford and C. Greenhalgh; "Increasing the Coherence between Human Beings and Virtual Agents",. Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems. Bologna. Italy, part 1, pp. 354-355, 2002.
- [14] P. Herrero, A. De Antonio; A Human Based Perception Model for Cooperative Intelligent Virtual Agents. Proceedings of the Tenth International Conference on Cooperative Information Systems (CoopIS 2002). Irvine. California. USA, pp. 195-212, 2002.
- [15] P. Herrero and A. De Antonio; "Introducing Human-like Hearing Perception in Intelligent Virtual Agents", To be published as full paper in proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems. Melbourne. Australia, 2003.
- [16] R. Hill; Soar: "An Architecture for Human Behavior Representation". Invited Panel Member and Presentation at the Workshop on Human Behavioral Representation, Conference on Modeling and Simulation. American Institute for Aeronautics and Astronautics (AIAA). Portland. Oregon, 1999.
- [17] R. Hill; "Modeling Perceptual Attention in Virtual Humans". Proceedings of the 8th Conference on Computer Generated Forces and Behavioral Representation, Orlando. Florida, 1999.
- [18] R. Hill, C. Han, M. Van Lent; "Perceptually Driven Cognitive Mapping of Urban Environments", Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems. Bologna. Italy, part 3, pp. 1389-1390, 2002.
- [19] R. Hill, Y.J. Kim and J. Gratch; "Anticipating Where to Look: Predicting the Movements of Mobile Agents in Complex Terrain", Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems. Bologna. Italy, part 2, pp. 821-827, 2002.
- [20] R. Hill, C. Han, M. Van Lent; "Applying Perceptually Driven Cognitive Mapping To Virtual Urban Environments; Conference on Innovative Applications of Artificial Intelligence (IAAI-2002). Edmonton. Canada, pp. 886-893, 2002.
- [21] N. Nilsson; "Teleo-Reactive Programs and the Triple-Tower Architecture", Electronic Transactions on Artificial Intelligence, Vol. 5 Section B, pp. 99-110, 2001.
- [22] H. Noser; "A Behavioral Animation System Based on L-systems and Synthetic Sensors for Actors", Ph.D. Thesis. École Polytechnique Fédérale De Lausanne. 1997.
- [23] J. Rickel, J. Gratch, R. Hill, S. Marsella and W. Swartout; "Steve Goes to Bosnia: Towards a New Generation of Virtual Humans for Interactive Experiences", Proceedings of AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment. Stanford University. California, pp. 67-71, 2001.
- [24] T. Rodden; Populating the Application; "A Model of Awareness for Cooperative Applications", Proceedings of ACM 1996 Conference on Computer Supported Co-operative Work (CSCW'96). Boston. Massachusetts. USA. ACM Press, pp. 87-96, 1996.
- [25] D. Terzopoulos, X. Tu and R. Grzeszczuk; Artificial fishes: "Autonomous locomotion, perception, behavior, and learning in a simulated physical world", Journal of Artificial Life, 1, 4, pp. 327-351, 1994.
- [26] D. Terzopoulos, T. F. Rabie; "Animat Vision: Active Vision with Artificial Animals; Proceedings of the Fifth International Conference on Computer Vision (ICCV'95), Cambridge. MA. USA, pp. 801-808, 1995.
- [27] D. Terzopoulos and T. F. Rabie; Animat Vision: "Active Vision in Artificial Animals", Videre: Journal of Computer Vision Research, 1(1), pp. 2-19, 1997.

- [28] D. Terzopoulos; "Biological and Evolutionary Models give life to self-animating graphical characters with bodies, brains, behavior, perception, learning, and cognition", *Artificial Life for Computer Graphics. Communications of the ACM*, 42 (8), 1999.  
<http://mrl.nyu.edu/~dt/papers/cacm99/cacm99.pdf>
- [29] D. Thalmann; "A New Generation of Synthetic Actors: the Interactive Perceptive Actors", *Proceedings of Pacific Graphics' 96. Taipeh. Taiwan*, pp. 200-219, 1996.
- [30] D. Thalmann; "The Foundations to Build a Virtual Human Society", *Proceedings of the 3rd International Workshop on Intelligent Virtual Agents (IVA'01)*. A. de Antonio, R. Aylett and D. Ballin (Eds.). Springer Lecture Notes in Artificial Intelligence, LNAI 2190. Madrid. Spain, pp. 1-14, 2001.
- [31] X. Tu and D. Terzopoulos; "Artificial Fishes: Physics, Locomotion, Perception, Behaviour", *Proceedings of Computer Graphics, SIGGRAPH 94*, pp 43-50, 1994.
- [32] W. K. Yeap and M. E. Jefferies; "Computing a representation of the local environment", *Artificial Intelligence* 107, pp. 265-301, 1999.
- [33] W. Zhang and R. Hill; "A Template-Based and Pattern-Driven Approach to Situation Awareness and Assessment in Virtual Humans", *Proceedings of the Fourth International Conference on Autonomous Agents, Barcelona, Cataluña, Spain*, ACM Press, pp. 116-123, 2000.