

NEUROFACES: AN INTELLIGENT SYSTEM FOR EMOTIONAL DESIGN OF AVATARS

Diego Más, José Antonio; Alcaide-Marzal, Jorge; Poveda Bautista, Rocío
Universitat Politècnica de València

Avatars are used in different contexts and situations: e-commerce, e-therapy, virtual worlds, videogames, collaborative online design... In this context, a good design of an avatar may improve the user experience. The ability of controlling the way an avatar convey messages and emotions is capital. In this work, a procedure to design avatar faces capable of conveying to the observer the most suitable sensations according to a given context is developed. The proposed system is based on a combination of genetic algorithms and artificial neural networks whose training is based on perceptual human responses to a set of faces.

Keywords: Human machine interaction; avatars; emotional design

NEUROFACES: UN SISTEMA INTELIGENTE PARA EL DISEÑO EMOCIONAL DE INTERLOCUTORES VIRTUALES

Los avatares o actores virtuales son utilizados en diferentes contextos y situaciones: e-commerce, e-terapia, mundos virtuales, videojuegos, diseño colaborativo en línea... En este contexto, un buen diseño de un avatar puede mejorar la experiencia del usuario. La capacidad de controlar la forma en que un avatar es capaz de transmitir mensajes y emociones es fundamental. En este trabajo, se desarrolla un procedimiento para diseñar caras de avatares o actores virtuales capaces de transmitir al observador la sensación más conveniente según un contexto determinado. El sistema propuesto se basa en una combinación de algoritmos genéticos y redes neuronales artificiales cuya formación se basa en respuestas perceptivas humanas a un conjunto de caras.

Palabras clave: Interacción hombre máquina; avatares; diseño emocional

1. Introducción

Internet se ha convertido en un lugar en el que las personas participan en muchas clases de interacciones, desde actividades profesionales (e-commerce, reuniones y entrevistas en línea, e-learning y educación...) a interacciones sociales (juegos, citas, comunidades...). Muchas de estas interacciones tienen lugar en un escenario virtual, generalmente llamado mundo virtual. Los mundos virtuales son entornos simulados con semejanza a los entornos reales en los que los participantes, mediante actores virtuales animados, pueden interactuar a través de herramientas generadas por ordenador (Bainbridge, 2007). El crecimiento de los sistemas de interacción online en todo tipo de actividades en internet acentúa la necesidad de personajes que representen a los usuarios y que les permita interactuar. En este contexto surgió el concepto de "avatar". Los avatares son definidos como "representaciones gráficas generales que son personificadas por medio de tecnología informática" (Holzwarth, Janiszewski y Neumann, 2006). Más específicamente, un avatar se define como una representa digital que simboliza la presencia del usuario en un metaverso (Davis, Murphy, Owens, Khazanichi & Zigurs, 2009).

Los avatares son utilizados en diferentes contextos y situaciones, y su morfología puede variar desde una pequeña imagen en un chat a un personaje interactivo en 3D en un videojuego o mundo virtual (Lin y Wang, 2014). Los avatares pueden representar personajes humanos, seres fantásticos, animales u objetos. Puede tener caracteres preconfigurados o personalizados por los usuarios. Sus características pueden modificarse y adaptarse de muchas maneras para adquirir una apariencia adecuada para un escenario dado.

Los usuarios persiguen que sus avatares comuniquen cosas sobre sí mismos, y emplean tiempo en elegirlos o personalizarlos. Se han llevado a cabo multitud de investigaciones sobre las estrategias de diseño de avatares para diversos propósitos (Kafai, Campos & Cook, 2010). Estos estudios muestran un equilibrio entre la necesidad de alinear la imagen de avatar con la del usuario y las estrategias de idealización. (Vasalou & Joinson, 2009), por ejemplo, muestra que los avatares diseñados para blogs reflejaban con precisión la apariencia física y personalidad de sus creadores. Sin embargo, cuando el avatar está destinado a ser utilizado en juegos de azar o citas, los usuarios destacan algunas características para que se ajusten mejor al contexto.

Por otra parte, la actitud y el comportamiento de las personas en los medios online pueden estar influidos por las características de su avatar. Varios estudios se han realizado en este sentido. En (Yee y Bailenson, 2007) se demuestra que la propensión de las personas a acercarse a otros y su auto-confianza, al establecer relaciones en línea, se ven afectados por la atracción y la altura de sus avatares. Resultados similares fueron observados en (Banakou et al., 2009). El trabajo que se muestra en (Merola & Peña, 2010) explica este efecto por confirmación conductual y argumenta que la interacción y la confianza online depende del uso de los avatares adecuados.

Los avatares son también empleados para transmitir emociones a través de expresiones no verbales. Se ha demostrado que el uso de emoticonos y avatares expresivos mejora las interacciones online (M Fabri & Moore, 2005) y puede incluso ser útil para la educación y con fines terapéuticos (Orvalho, Miranda & Sousa, 2009). Además, la transmisión de las emociones está fuertemente relacionado con la credibilidad, que a su vez afecta a la capacidad de persuadir (El-Nasr, Ioerger, Yen, House y Parke, 1999). Esta capacidad es importante para aplicaciones en áreas que van desde el e-commerce a la e-terapia. Las características de un avatar en un sitio web de comercio electrónico pueden mejorar la experiencia del usuario e influir en la generación de confianza (Keeling, McGoldrick & Beatty, 2010). En (Holzwarth et al., 2006) se muestra que un avatar atractivo es más eficaz en la venta de productos online. El grupo étnico y género del avatar también afectan a su

percepción como agente de venta (Qiu, 2006). La importancia de los avatares en el marketing online es tal que algunos autores sostienen que las empresas deben desarrollar estrategias específicas centradas en las características de los avatares que emplean en sus sitios web (Hemp, 2006). Por lo que se muchos estudios se centran en lograr mejorar la expresividad y realismo de los avatares (Wang y Geigel, 2011).

Sin embargo, la transmisión de emociones a través de avatares depende en gran medida de la percepción y evaluación de estas expresiones por el espectador. Varios estudios han reportado diferencias culturales en la interpretación de expresiones faciales de avatares (; Tomoko Koda, Ishida, Rehm y André, 2009). Algunos otros trabajos exploran maneras de desarrollar avatares emocionales más creíbles. En (El-Nasr et al., 1999) se propone un algoritmo para generar avatares 2D con expresiones emocionalmente dinámicas. En (Mazzei, Lazzeri, Hanson & De-Rossi, 2012) se propone un software llamado HEFES para crear expresiones faciales en un robot y su avatar 3D y emplearlos en el tratamiento de niños con autismo. Otros sistemas para generar gestos expresivos en avatares 3D pueden encontrarse en (Yu, Garrod y Schyns, 2012).

En este contexto, la capacidad de controlar la manera en la que un avatar transmite al observador mensajes y emociones es capital. El propósito de esta investigación es desarrollar un procedimiento para el diseño de caras de avatares capaces de transmitir al observador la sensación más conveniente según un contexto determinado, generalmente en aplicaciones informáticas. En este contexto llamamos "sensación" a la impresión que las cosas producen por medio de los sentidos y que se produce inmediatamente después de su excitación. Es una experiencia inmediata básica generada por estímulos aislados simples (Foley y Matlin, 2010). Evidentemente, la observación de un avatar producirá sensaciones en los observadores, el problema es que existe una elevada dispersión en el juicio de los observadores sobre las sensaciones producidas. Aunque existe una relación entre los atributos de la cara del avatar y las sensaciones que produce, varía de un observador a otro debido a factores individuales y externos (motivaciones, educación, personalidad...) (Hoch, Kim, Montgomery, Rossi, & Montgomery, 1995).

Dado que un mismo avatar será utilizado por muchas personas, el objetivo será transmitir la sensación deseada a la mayoría de ellos. El sistema propuesto se basa en una combinación de algoritmos genéticos (GAs) y redes neuronales artificiales (ANNs) cuyo entrenamiento se basa en las sensaciones generadas a un grupo numeroso de personas por un conjunto de amplio de caras de avatares.

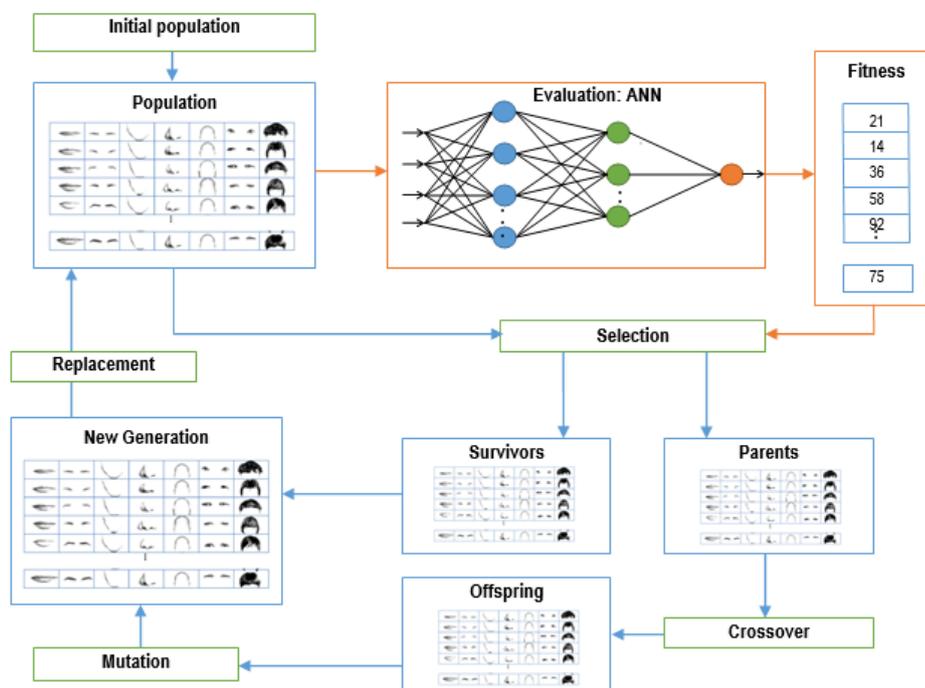
2. Metodología

Como se ha indicado en el apartado 1, resulta importante en el proceso de interacción usuario-avatar conseguir que el aspecto de este último transmita al usuario determinadas sensaciones. Para ello es necesario determinar la combinación de rasgos que, en opinión de los usuarios potenciales, transmite de forma óptima dichas sensaciones. La gran diversidad de los rasgos faciales provoca que el número de avatares que es posible crear a partir de la combinación de cierto número de ellos sea enorme. Por ello resulta inviable hacer una evaluación de la sensación que transmiten a los usuarios todos los posibles avatares para obtener el avatar óptimo. El empleo de GAs resulta adecuado cuando no es posible obtener la solución del problema de forma analítica, y cuando el tamaño del espacio de soluciones es demasiado grande como para afrontar la búsqueda de soluciones de una manera enumerativa.

Los GAs realizan una búsqueda estocástica guiada basada en la repetición reiterada de una serie de estructuras denominadas cromosomas. El punto de comienzo es un conjunto de soluciones del problema denominado población inicial (Figura 1). En el presente caso, un individuo o solución será la cara de un avatar formada por la combinación de diferentes

rasgos codificados en un cromosoma. Cada individuo es evaluado por medio de una función objetivo, en este caso, en qué grado es capaz de provocar en el observador la sensación que se persigue. La población sufre varias transformaciones que producen una nueva población (nueva generación). Estas transformaciones son guiadas por algunos operadores genéticos, los más comunes son la selección, el cruce y la mutación, que combinan o modifican los cromosomas de los individuos. Los operadores de cruce y mutación se aplican para crear una nueva generación de individuos que heredan las mejores características de sus predecesores. Para este propósito, los individuos que participan en cada uno de los operadores genéticos y los que sobreviven y pasan a la siguiente generación son seleccionados previamente por medio del operador de selección. Se repite el proceso con el nuevo conjunto de individuos hasta llega a un cierto número de iteraciones o hasta que se ha realizado un cierto número de iteraciones sin una nueva solución mejor, haciendo que los individuos evolucionen a mejores soluciones al problema. Como se ha indicado, los individuos deben ser evaluados para establecer en qué medida son una buena solución al problema que se trata de resolver. De forma habitual la evaluación se realiza mediante fórmulas analíticas de cálculo, pero en este caso emplear este procedimiento no es posible porque no existe una fórmula que permita determinar si un determinado avatar es capaz de provocar en el usuario determinada sensación. Por otra parte, si el tamaño de la población empleado por el GA es, por ejemplo, 50, en 1000 generaciones se pueden generar 50000 avatares diferentes, por lo que no es operativo emplear a la opinión de los usuarios potenciales como función evaluadora. Para solventar este problema puede emplearse una ANN entrenada para predecir las opiniones que los usuarios tendrán sobre cada avatar.

Figure 1. Estructura del GA empleando una ANN como función evaluadora.



Una ANN es un modelo matemático que representa un sistema adaptativo distribuido construido con múltiples elementos de procesamiento interconectados, como las redes neuronales reales. Las ANNs son utilizadas en muchos campos de la investigación (ciencia del diseño, psicología, robótica, biología, producción o computadora, para nombrar unos pocos) (Principe, Euliano y Lefebvre, 2000) debido a su capacidad de adaptarse, aprender, generalizar, organizar o agrupar datos. Debido a su capacidad para aprender (en

comparación con los sistemas secuenciales), son instrumentos adecuados muchos usos (Chen & Yan, 2008; Dasgupta, Dispensa y Ghose, 1994; Hsiao y Tsai, 2005; Ishihara, Ishihara, Nagamachi y Matsubara, 1997; Lai, Lin y Yeh, 2005; Lai, Lin, Yeh y Wei, 2006; Tsai, Hsiao y Hung, 2006; Yang y Shieh, 2010).

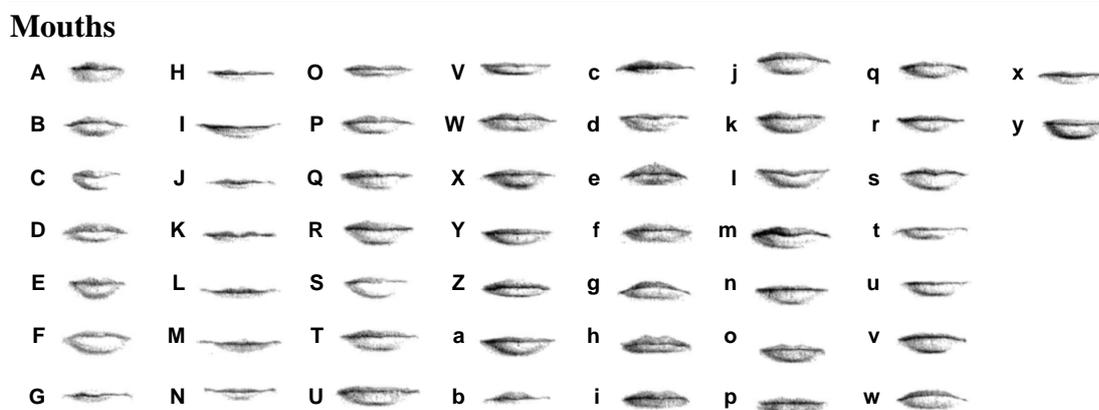
Los elementos de procesamiento (neuronas) están distribuidos en varias capas. Las capas intermedias son conocidas como capas ocultas, mientras que la primera y última capas se conocen como la capa de entrada y de salida respectivamente. En términos generales, cada neurona recibe señales, las procesa y las transmite. El número de capas y la forma en que están conectadas las neuronas determina la arquitectura de la red.

Así pues, el procedimiento propuesto en este trabajo consiste en generar una ANN entrenada para juzgar los avatares a partir de las opiniones de los usuarios potenciales de los mismos. Dicha ANN se empleará como función objetivo de un GA que determine la combinación de atributos del avatar que mejor provoque la sensación deseada en los usuarios potenciales. Para comprobar el funcionamiento de este procedimiento se plantea a continuación un caso de estudio.

3. Caso de estudio

Como caso de estudio se decidió emplear el procedimiento descrito en la sección anterior para generar las caras de tres avatares masculinos cuyo aspecto fuera el de personas caucásicas de entre 20 y 30 años de edad. Para comprobar que es posible generar de esta forma avatares que inspiren diferentes sensaciones en el observador, el primero debería ser considerado como atractivo, el segundo una persona que inspira confianza y el tercero una persona responsable. Se consideró que los usuarios potenciales pertenecían al mismo rango de edad que los avatares (entre 20 y 30 años).

Figura 2. Ejemplo de la base de datos de rasgos. Cada rasgo está codificado con una letra.



Para la generación de las caras se decidió emplear siete rasgos faciales masculinos: boca, cejas, mandíbula, nariz, cabeza, ojos y pelo. Se generó una base de datos de imágenes de rasgos faciales formada por 51 bocas, 32 cejas, 29 mandíbulas, 38 narices, 9 tipos de cabeza, 22 pares de ojos y 52 tipos de pelo (Figura 2). El número de caras que es posible generar con la combinación de estos rasgos, y por tanto el espacio de soluciones a explorar por el GA es, aproximadamente, $1,852e+10$.

Se empleó Embarcadero RAD Studio™ para desarrollar un software de apoyo al proceso completo de generación de los avatares. Este software permitía generar caras de avatares

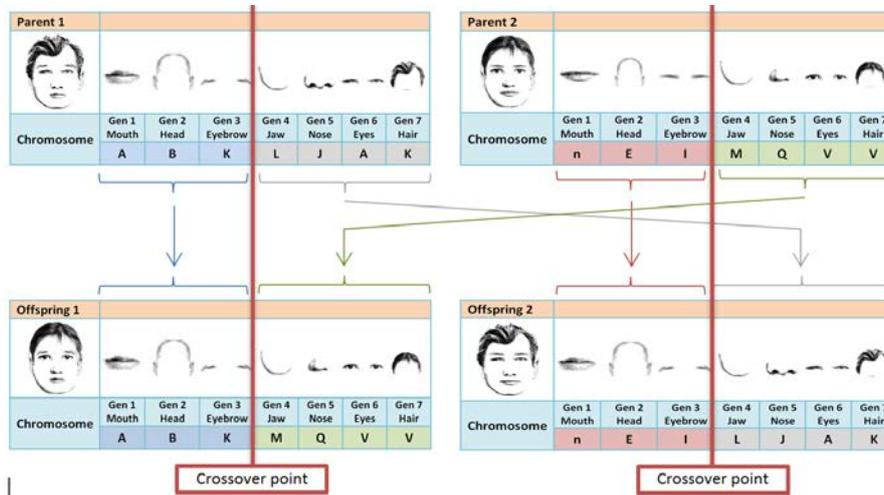
combinando los diferentes rasgos de la base de datos, obtener datos para el entrenamiento de una ANN y emplear un GA para generar los avatares a partir de la ANN entrenada.

3.1 Desarrollo del algoritmo genético

El algoritmo comienza por generar una población inicial de individuos que representan diferentes avatares (Figure 1). Cada individuo es codificado a través de un vector (cromosoma). La codificación empleada consistió en un vector de longitud siete en el que cada elemento es un gen que representa uno de los rasgos empleados para generar la cara del avatar (Figure 3). Cada gen representa, en este orden, la boca, la forma de la cabeza, las cejas, la mandíbula, la nariz y el pelo del avatar. A cada rasgo se le asignó el código alfabético mostrado en la Figura 2 (diferenciando entre letras mayúsculas y minúsculas). Así pues, la codificación de una cara determinada se lleva a cabo indicando el código de cada rasgo en el gen correspondiente. Para formar los cromosomas de la población inicial se eligieron aleatoriamente los valores de los genes.

Cada uno de los rostros que conforman la población inicial debe ser evaluado para determinar en qué medida dicho conjunto de rasgos forman una cara capaz de provocar en el usuario la sensación deseada. Como se ha dicho, este proceso será llevado a cabo mediante una ANN con una única salida. Las entradas serán los códigos de cada uno de los genes y la salida la valoración de la cara correspondiente. Dado que las entradas a la red son variables categóricas, los siete tipos rasgos serán codificados de forma unaria. Cada tipo de rasgo es desglosado en tantas neuronas como rasgos diferentes existen. Por ejemplo, el rasgo Head se desglosará en 9 neuronas puesto que existen 9 tipos de cabezas. La cabeza A se representará con un 1 en la entrada de la primera neurona y un cero en las restantes 8. El modo de obtener la ANN será expuesto en la Sección 3.2.

Figura 3. Operador cruce



Cuando todos los individuos han sido evaluados, se realiza una selección de los que sobrevivirán y pasarán a la siguiente generación o se utilizarán como reproductores. Para este proceso que se utiliza una selección de ruleta (Goldberg, 1989), en la que la probabilidad de que un individuo sea seleccionado está relacionado con el valor obtenido en su evaluación, de tal forma que aquellos individuos que ofrecen los mejores resultados tienen mayor probabilidad de ser seleccionados. Pares de individuos son elegidos entre los seleccionados en la etapa anterior. La reproducción se realiza por medio de un cruce, en el cual se generan dos nuevos individuos a partir de la combinación de las soluciones representadas por los individuos originales (véase la Figura 3). El parámetro P_c (probabilidad de cruce) determina el número de individuos en la generación que será creada

por el cruce y el número de individuos que serán los supervivientes. Un valor típico de este parámetro varía entre 0.5 y 0.9 (Srinivas & Patnaik, 1994). El punto de cruce es elegido al azar. Los descendientes se generan mediante la combinación de los cromosomas que quedan a la izquierda y a la derecha del punto de cruce en cada uno de los padres. El operador de mutación se aplica a los individuos seleccionados al azar. El parámetro que determina el número de individuos que se mutan es P_m (probabilidad de mutación). El proceso de mutación consistió en elegir de manera aleatoria uno de los genes del cromosoma, cambiando su valor por otro de los posibles valores de ese rasgo. Los individuos generados por cruce y los supervivientes conforman una nueva población que sustituye a la anterior. El proceso se repitió con cada nuevo conjunto de individuos hasta que se llegó a un cierto número de iteraciones.

3.2 –Obtención de las redes neuronales

Para el desarrollo de la ANN que debe actuar como función evaluadora del GA es necesario obtener un número suficiente de patrones del proceso que se quiere modelizar. En este caso, un patrón consiste en un cromosoma representando una cara y la respuesta correspondiente dada por un usuario real. Estos patrones serán empleados en el proceso de determinación del tipo de ANN más adecuado, de la topología de la red (número de capas, número de neuronas, funciones de transferencia...) y para entrenar a la red escogida. Dado que en este caso de estudio se valoraron tres sensaciones diferentes fue necesario obtener tres ANNs.

3.2.1 – Obtención de datos para el entrenamiento de las ANNs

26 personas de entre 20 y 30 años fueron seleccionadas para ser encuestadas (15 men and 11 women). Para llevar a cabo la encuesta el software desarrollado presentó a cada usuario 40 caras generadas aleatoriamente empleando los rasgos de la Figura 2 junto con el juicio que se espera del observador. Cada usuario debía valorar cada cara en una escala de 0 a 100, siendo 0 “Completamente de acuerdo” and 100 “en completo desacuerdo”. Este procedimiento se repitió tres veces con las diferentes sensaciones que se querían medir: “Considera atractiva a esta persona?”, “Considera que esta persona inspira confianza?” y “Considera que ésta es una persona responsable?”. Para cada una de las tres sensaciones se obtuvieron 1040 casos. Cada caso estaba formado por los rasgos empleados para obtener la cara del avatar y la correspondiente valoración del usuario. Los casos de cada sensación fueron divididos tres sets de datos. 790 casos formaron el training set, empleado para entrenar la red. Para crear el set de validación se emplearon 226 casos. Los restantes 24 casos formaron el test set.

3.2.2 –Entrenamiento de las redes

Se utilizó NeuroSolution 5.0 para crear y entrenar a las ANNs. Tras algunas pruebas se escogió una Multilayer Perceptron como el tipo de red más adecuado para este problema, y el algoritmo de entrenamiento utilizado fue back propagation con momentum. El momentum es un término que introduce un cierto grado de inercia en la actualización de los pesos de las conexiones neuronales, acelerando y estabilizando la evolución de los pesos y evitando la convergencia en un mínimo local. El momentum se fijó en 0,7. El tamaño de paso utilizado fue 0.01 en la primera capa oculta, reducida por un factor de 5 en la segunda capa oculta y nuevamente en la tercera (Principe et al., 2000). Los pesos iniciales fueron establecidos al azar con valores pequeños (con media 0 y rango entre -0.1 y 0.1). Este procedimiento se recomienda cuando se utilizan funciones de activación no lineales para evitar la saturación inicial de los elementos de procesamiento (Principe et al., 2000). Se utilizaron como funciones de activación tangentes hiperbólicas.

Después de varias pruebas se seleccionaron las mejores redes obtenidas para cada sensación, empleando como criterio el mean square error (MSE). Las características de las

tres mejores ANNs encontradas se muestran en la Tabla 1. Las valoraciones para las que se entrenaron las redes fueron “¿Considera atractiva a esta persona?”, “¿Considera que esta persona inspira confianza?” y “¿Considera que ésta es una persona responsable?” respectivamente. Como ejemplo, la Figura 4 muestra las opiniones de los usuarios sobre las 24 caras del test set para la pregunta “¿Considera atractiva a esta persona?”, y los resultados obtenidos por la ANN1.

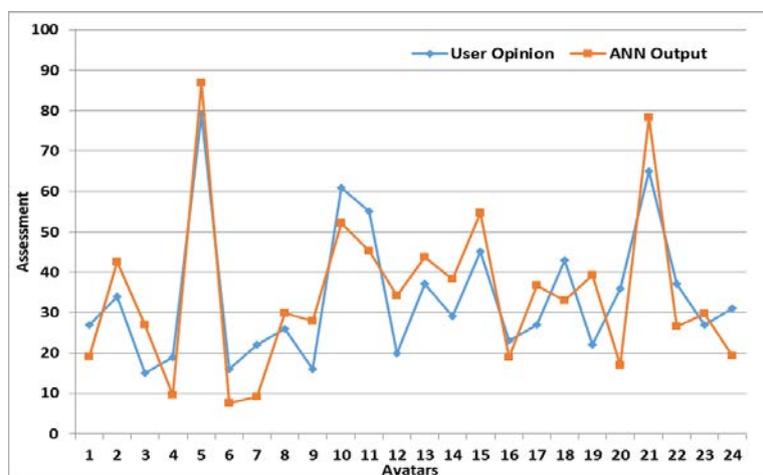
4 - Resultados

Una vez obtenidas las ANNs se empleó el software desarrollado para la ejecución del GA descrito en la sección 3.1. Después de algunas pruebas para escoger los parámetros de ejecución más adecuados, se emplearon poblaciones de 50 individuos y el máximo número de generaciones fue fijado en 1000. El GA fue ejecutado tres veces, una con cada ANN obtenida en la sección anterior. El experimento fue repetido otras tres veces empleando los mismos parámetros pero siendo la función objetivo minimizar las salidas de las ANNs. La Figura 5 muestra los cromosomas que obtuvieron las puntuaciones máximas y mínimas en cada caso, así como el fenotipo correspondiente a cada uno de ellos.

Tabla 1. Características de las mejores ANNs localizadas

	ANN 1	ANN 2	ANN 3
First Layer Neurons	8	6	9
Second Layer Neurons	4	3	4
Test MSE	0.0367	0.0209	0.0706

Figura 4. Comparación de las opiniones de los usuarios y las predichas por la ANN1.



Para comprobar los resultados obtenidos por el GA con la opinión de usuarios reales, se pidió a 9 usuarios que, empleando el software desarrollado, combinaran los rasgos de la base de datos para obtener la cara que consideraran que transmitía mejor cada una de las sensaciones. Se dio a los usuarios todo el tiempo que necesitaran para obtener la combinación más conveniente. De esta forma se obtuvieron 9 caras bien valoradas por los usuarios para cada tipo de sensación a las que se añadieron las obtenidas por el GA. Los tres sets de 10 caras fueron presentados a 10 usuarios. Se les pidió que las ordenaran en función del grado en que transmitían cada sensación, siendo la primera la que lo hacía en mayor medida y la última la que menos lograba transmitir la sensación. Una vez ordenadas se calculó la posición promedio de cada cara.

La Tabla 2 muestra la posición promedio de las caras obtenidas por el GA en los tres tipos de sensaciones. Como puede observarse, aunque no en todos los casos las caras

desarrolladas fueron las mejor valoradas por los usuarios, sí obtuvieron la mejor posición promedio, 1.3, 1.5 y 1.7 respectivamente, siendo las que más veces se clasificaron en 1ª posición. Las caras con la segunda mejor posición promedio en cada caso obtuvieron 2.4, 2.8 y 2.9. Por otra parte, la dispersión de opiniones fue menor en las caras obtenidas con el GA en los tres casos.

Figura 5. Avatares con la mejor y peor puntuación para cada sensación localizados por el GA.

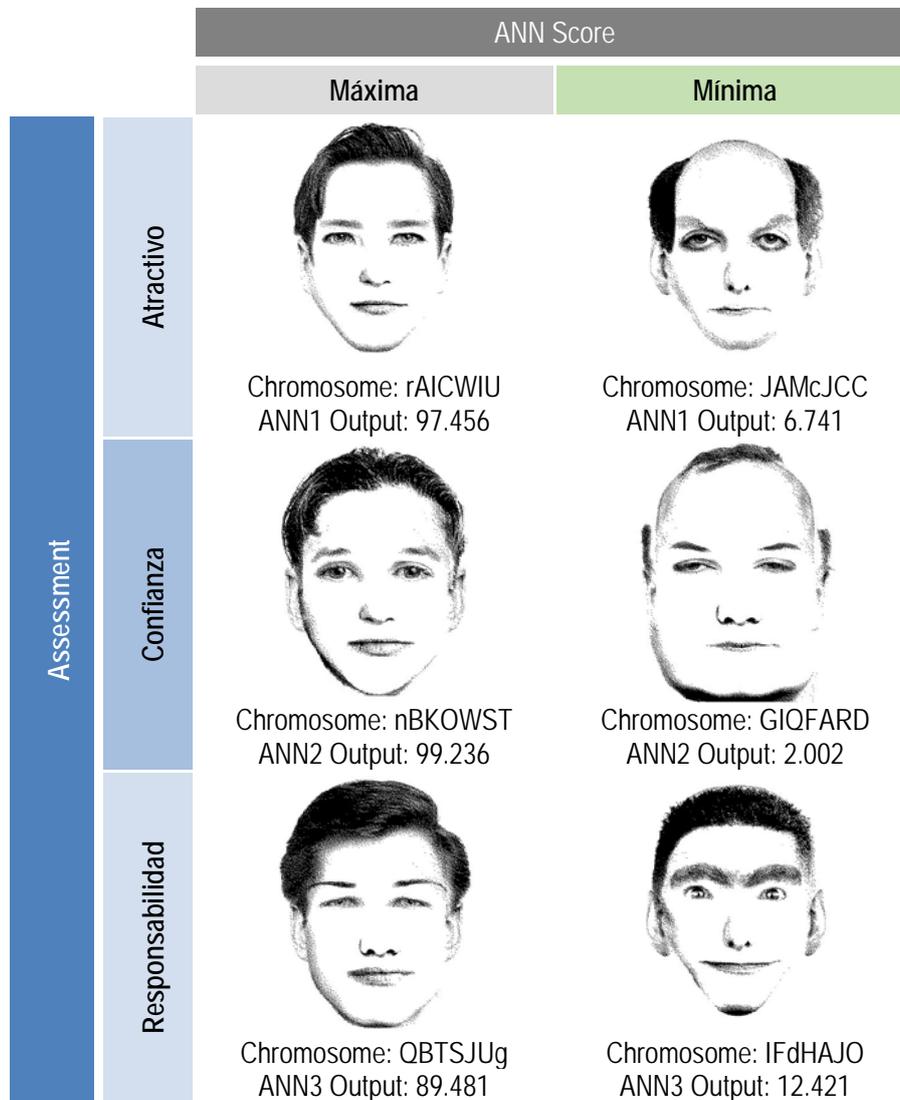


Tabla 2. Posición promedio de las caras obtenidas por el GA en los tres tipos de sensaciones según usuarios reales.

	Posición promedio	Veces en 1ª posición	Veces en 2ª posición	Veces en 3ª posición
Atractivo	1.3	7	3	0
Confianza	1.5	6	3	1
Responsabilidad	1.7	4	5	1

5. Discusión y conclusiones.

Los resultados obtenidos muestran que el proceso propuesto es capaz de generar caras de avatares que transmiten a los usuarios la sensación deseada. El procedimiento ha logrado dicho objetivo con tres tipos de sensaciones diferentes: atractivo, confianza y responsabilidad. Los mejores avatares encontrados fueron, en promedio, los mejor valorados para cada sensación por los usuarios. No todos los usuarios consideraron estos avatares como los más adecuados, pero cabe recordar que el proceso propuesto se basa en las opiniones de muchos usuarios distintos, por lo que el procedimiento obtiene avatares que maximizan la opinión media de los usuarios.

La Figura 4 muestra las diferencias entre las opiniones de los usuarios y las obtenidas por la ANN1 respecto a las caras del test set. Puede observarse que, aunque esto no ha supuesto un problema importante para el procedimiento descrito, en algunos casos existen diferencias significativas entre las opiniones del usuario real y la valoración de la red. La ANN ha sido entrenada con caras valoradas por un único usuario. Como se ha mencionado existe una importante variabilidad en las opiniones de los usuarios debido a gustos particulares, por lo que, en general, la valoración obtenida por la red no tiene por qué coincidir con la de un usuario individual. Esto es debido a que, durante el entrenamiento, la red ha adaptado los pesos de sus conexiones para minimizar el error entre su salida y las opiniones de muchos usuarios diferentes. Una alternativa a este procedimiento de entrenamiento es que cada cara empleada para entrenar a las redes sea valorada por varios usuarios, empleando la valoración media obtenida por cada una. Este procedimiento es más costoso puesto que es necesario que cada usuario valore un número mayor de caras para entrenar las redes. Todas las caras deberían ser valoradas por todos los usuarios, y dado que es necesario que exista un número suficiente de casos para entrenar las redes, el tiempo empleado por los usuarios en la encuesta sería demasiado grande.

Por otra parte, el MSE sobre los casos de test de las ANNs ha sido muy diferente para cada sensación. El MSE de la sensación Responsabilidad (0.0706) es más del triple que el de la sensación Confianza (0.0209). Esta gran diferencia puede deberse a diversos motivos. Es posible que la dispersión de las opiniones de los usuarios sea mayor en unas sensaciones que en otras, lo que dificulta el ajuste de la red. Puede ocurrir, por otra parte, que las palabras empleadas para describir las sensaciones tengan significados menos definidos en unos casos que en otros, es decir, que, por ejemplo, el consenso sobre el significado de la palabra Confianza sea mayor que el del significado de la palabra Responsable.

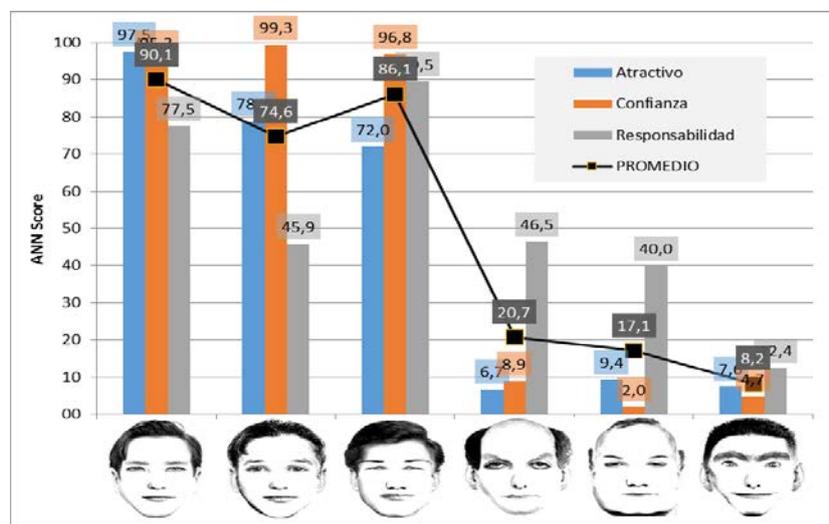
Además de obtener caras de avatares que transmitan la sensación deseada para un determinado contexto, minimizando la función objetivo del algoritmo genético (la salida de la ANN) es posible obtener avatares que transmitan la sensación deseada lo menos posible. Es necesario aclarar que las ANN fueron entrenadas para medir una sensación determinada, y que al minimizar la salida de las ANN no se están obteniendo necesariamente avatares que transmiten la sensación contraria. Por ejemplo, puede decirse que la sensación contraria a Confianza es Desconfianza, pero la ANN2 fue entrenada para medir la Confianza y no la Desconfianza. Una cara con una valoración baja de la red transmitirá poca Confianza, pero no necesariamente poca confianza equivale a la sensación de Desconfianza.

Los resultados han mostrado que las caras obtenidas con este procedimiento son mejor valoradas de forma global que las creadas por los propios usuarios. Además de obtener las mejores posiciones promedio al clasificar un conjunto de caras, la dispersión de las opiniones respecto a ellas es menor. Por ejemplo, la cara obtenida por el GA para la sensación Atractivo obtuvo una posición promedio de 1.3 para 10 usuarios ocupando en 7 casos la primera posición y en 3 la segunda. La segunda mejor cara obtuvo una posición promedio de 2.4, y la dispersión de las opiniones fue mayor: la primera en 2 casos, la

segunda en 4, la tercera en 2 y la cuarta en 2. Esto ocurrió de la misma forma con las tres sensaciones valoradas. Así pues, podría concluirse que el procedimiento propuesto obtiene caras con la mejor valoración promedio y con un mayor consenso en la opinión de los usuarios.

Aunque los resultados obtenidos son satisfactorios, es necesario investigar más profundamente algunos aspectos. Se ha empleado caras de avatares formadas por siete tipos de rasgos y un total de 233 rasgos diferentes. Sin embargo, es muy probable que los resultados obtenidos puedan mejorar aumentando el número de tipos de rasgos y la cantidad de rasgos de cada tipo. Podrían considerarse aspectos como el color de los ojos, los dientes o el color de pelo. Por otra parte, podría incluirse la posición relativa de los diferentes rasgos en la cara del avatar. Además, el procedimiento desarrollado puede extenderse al avatar completo, incluyendo ropa, accesorios de diferentes tipos y complejiones corporales distintas.

Figura 6. Valoración de las seis caras obtenidas en el caso de estudio para cada una de las tres sensaciones medidas.



Por último, es posible emplear este procedimiento para la generación de avatares que transmitan varias sensaciones simultáneamente. Por ejemplo, podría ser deseable obtener un avatar que pareciera atractivo, confiable y responsable de forma simultánea. Una vez desarrolladas las correspondientes ANNs para cada sensación, puede emplearse como función objetivo del GA una suma ponderada de las salidas de dichas redes. Como ejemplo, la Figura 6 muestra la valoración de las seis caras obtenidas en el caso de estudio para cada una de las tres sensaciones medidas empleando las 3 ANNs simultáneamente, así como la valoración promedio de las tres sensaciones.

Bibliografía

- Bainbridge, W. S. (2007). The scientific research potential of virtual worlds. *Science* (New York, N.Y.), 317(5837), 472–476. doi:10.1126/science.1146930
- Brace, I. (2013). Questionnaire design: how to plan, structure, and write survey material for effective market research. (K. Page, Ed.) (3rd ed., p. 288). London.
- Chen, C., & Yan, W. (2008). An in-process customer utility prediction system for product conceptualisation. *Expert Systems with Applications*, 34(4), 2555–2567. doi:10.1016/j.eswa.2007.04.019

- Dasgupta, C. G., Dispensa, G. S., & Ghose, S. (1994). Comparing the predictive performance of a neural network model with some traditional market response models. *International Journal of Forecasting*. doi:10.1016/0169-2070(94)90004-3
- Davis, A., Murphy, J., Owens, D., Khazanchi, D., & Zigungs, I. (2009). Avatars, people, and virtual worlds: Foundations for research in metaverses. *Journal of the Association for Information Systems*, 10(2), 90–117. doi:1660426061
- Ducheneaut, N., Wen, M., Yee, N., & Wadley, G. (2009). Body and Mind: A Study of Avatar Personalization in Three Virtual Worlds. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1151–1160. doi:10.1145/1518701.1518877
- El-Nasr, M. S., Ioerger, T. R., Yen, J., House, D. H., & Parke, F. I. (1999). Emotionally expressive agents. *Proceedings Computer Animation 1999*. doi:10.1109/CA.1999.781198
- Fabri, M., & Elzouki, S. (2007). Emotionally expressive avatars for chatting, learning and therapeutic intervention. In J. A. Jacko (Ed.), *Human-Computer Interaction, HCI Intelligent Multimodal Interaction Environments 12th International Conference* (pp. 275–285). Springer Berlin / Heidelberg. doi:10.1007/978-3-540-73110-8
- Fabri, M., & Moore, D. (2005). The use of emotionally expressive avatars in Collaborative Virtual Environments. *Sin Nombre*.
- Foley, H. J., & Matlin, M. W. (2010). *Sensation and perception* (5th ed.). Boston: Allyn & Bacon.
- Hasler, B. S., Tuchman, P., & Friedman, D. (2013). Virtual research assistants: Replacing human interviewers by automated avatars in virtual worlds. *Computers in Human Behavior*, 29, 1608–1616. doi:10.1016/j.chb.2013.01.004
- Hemp, P. (2006). Avatar-Based Marketing. *Harvard Business Review*, 84(6), 48–57. doi:10.1037/e661492007-001
- Hoch, S. J., Kim, B., Montgomery, A. L., Rossi, P. E., & Montgomery, L. (1995). Determinants of Store-Level Price. *Journal of Marketing Research*, 32, 17–29. doi:10.2307/3152107
- Holzwarth, M., Janiszewski, C., & Neumann, M. M. (2006). The Influence of Avatars on Online Consumer Shopping Behavior. *Journal of Marketing*. doi:10.1509/jmkg.70.4.19
- Hsiao, S.-W., & Huang, H. (2002). A neural network based approach for product form design. *Design Studies*, 23(1), 67–84. doi:10.1016/S0142-694X(01)00015-1
- Hsiao, S.-W., & Tsai, H.-C. (2005). Applying a hybrid approach based on fuzzy neural network and genetic algorithm to product form design. *International Journal of Industrial Ergonomics*, 35(5), 411–428. doi:10.1016/j.ergon.2004.10.007
- Ishihara, S., Ishihara, K., Nagamachi, M., & Matsubara, Y. (1997). An analysis of Kansei structure on shoes using self-organizing neural networks. *International Journal of Industrial Ergonomics*. doi:10.1016/S0169-8141(96)00006-6
- Kafai, Y. B., Fields, D. A., & Cook, M. S. (2010). Your Second Selves: Player-Designed Avatars. *Games and Culture*. doi:10.1177/1555412009351260
- Keeling, K., McGoldrick, P., & Beatty, S. (2010). Avatars as salespeople: Communication style, trust, and intentions. *Journal of Business Research*, 63(8), 793–800.
- Koda, T. (2004). Interpretation of Emotionally Expressive Characters in an Intercultural Communication. In M. Negoita, R. Howlett, & L. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems SE - 113* (Vol. 3214, pp. 862–868). Springer Berlin Heidelberg. doi:10.1007/978-3-540-30133-2_113
- Koda, T., Ishida, T., Rehm, M., & André, E. (2009). Avatar culture: cross-cultural evaluations of avatar facial expressions. *AI & SOCIETY*. doi:10.1007/s00146-009-0214-5
- Lai, H.-H., Lin, Y.-C., & Yeh, C.-H. (2005). Form design of product image using grey relational analysis and neural network models. *Computers & Operations Research*, 32(10), 2689–2711. doi:10.1016/j.cor.2004.03.021

- Lai, H.-H., Lin, Y.-C., Yeh, C.-H., & Wei, C.-H. (2006). User-oriented design for the optimal combination on product design. *International Journal of Production Economics*, 100(2), 253–267. doi:10.1016/j.ijpe.2004.11.005
- Lee, J.-E. R. (2014). Does virtual diversity matter?: Effects of avatar-based diversity representation on willingness to express offline racial identity and avatar customization. *Computers in Human Behavior*, 36, 190–197. doi:http://dx.doi.org/10.1016/j.chb.2014.03.040
- Lin, H., & Wang, H. (2014). Avatar creation in virtual worlds: Behaviors and motivations. *Computers in Human Behavior*, 34, 213–218.
- Mazzei, D., Lazzeri, N., Hanson, D., & De-Rossi, D. (2012). HEFES: An Hybrid Engine for Facial Expressions Synthesis to control human-like androids and avatars. *Biomedical Robotics and Biomechanics (BioRob)*, 2012 4th IEEE RAS & EMBS International Conference on, 195–200. doi:10.1109/BioRob.2012.6290687
- Merola, N., & Pena, J. (2010). The Effects of Avatar Appearance in Virtual Worlds. *Journal of Virtual Worlds Research*, 2(5), 3–12.
- Orvalho, V., Miranda, J., & Sousa, A. A. (2009). Facial Synthesis of 3D Avatars for Therapeutic Applications. *Studies in Health Technology and Informatics*, 144, 96–98.
- Principe, J. C., Euliano, N. R., & Lefebvre, W. . (2000). *Neural and Adaptive Systems: Fundamentals Through Simulations*. New York: Wiley and Sons.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. . Rumelhart & J. L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 1, pp. 318–362). Cambridge: The MIT Press.
- Sarle, W. S. (1995). Stopped Training and Other Remedies for Overfitting. *Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics*, 352–360.
- Savage, S. J., & Waldman, D. M. (2008). Learning and fatigue during choice experiments: a comparison of online and mail survey modes. *Journal of Applied Economics*, 23, 351–371. doi:10.1002/jae.984
- Srinivas, M., & Patnaik, L. M. (1994). Genetic algorithms: a survey. *Computer*, 27. doi:10.1109/2.294849
- Tsai, H.-C., Hsiao, S.-W., & Hung, F.-K. (2006). An image evaluation approach for parameter-based product form and color design. *Computer-Aided Design*, 38(2), 157–171. doi:10.1016/j.cad.2005.10.001
- Vasalou, A., Joinson, A., Bänziger, T., Goldie, P., & Pitt, J. (2008). Avatars in social media: Balancing accuracy, playfulness and embodied messages. *International Journal of Human Computer Studies*, 66(11), 801–811.
- Wang, Y., & Geigel, J. (2011). Using Facial Emotional Signals for Communication between Emotionally Expressive Avatars in Virtual Worlds. In 2011 Humaine Association Conference on Affective Computing and Intelligent Interaction (pp. 297–304). doi:10.1007/978-3-642-24571-8_33
- Yang, C.-C., & Shieh, M.-D. (2010). A support vector regression based prediction model of affective responses for product form design. *Computers & Industrial Engineering*, 59(4), 682–689. doi:10.1016/j.cie.2010.07.019
- Yee, N., & Bailenson, J. (2007). The proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, 33(3), 1–38. doi:10.1111/j.1468-2958.2007.00299.x
- Yu, H., Garrod, O. G. B., & Schyns, P. G. (2012). Perception-driven facial expression synthesis. *Computers & Graphics*, 36(3), 152–162. doi:10.1016/j.cag.2011.12.002