

Resumen

El campo de la computación afectiva es un área que ha surgido con gran impulso y está en constante evolución. Este campo logra integrar psicofisiología, informática, ingeniería biomédica e inteligencia artificial, desarrollando sistemas capaces de inducir y reconocer emociones de manera automática. Su enfoque principal es el estudio del comportamiento humano a través de las emociones, las cuales desempeñan un papel fundamental en acciones como la interacción social, la toma de decisiones o la memoria.

Recientemente, los avances tecnológicos han posibilitado el desarrollo de sistemas inteligentes humano-máquina que antes no eran factibles. Elementos como la realidad virtual y los modelos generativos de inteligencia artificial están adquiriendo un papel relevante en el campo de la computación afectiva. La combinación de estas tecnologías puede dar lugar a experimentaciones mucho más realistas que, junto con el reconocimiento automático de respuestas fisiológicas, podrían constituir una metodología robusta para evocar e identificar estados emocionales en entornos de dinámicas sociales.

Esta tesis se centra en la evocación y reconocimiento de emociones en realidad virtual mediante la creación del primer humano virtual, basado en un modelo de lenguaje generativo, que permite mantener conversaciones a tiempo real con un humano. Se han medido diversas señales fisiológicas y desarrollado herramientas de procesamiento de señales para monitorizar las respuestas de manera automática. A partir de esto, se ha evaluado y validado no solo la evocación y reconocimiento de emociones mediante aprendizaje automático, sino también el reconocimiento de sujetos con síntomas depresivos.

Este trabajo comienza con la adaptación de un algoritmo de fijaciones de *eye-tracking* en 2D a un entorno 3D de realidad virtual. En este contexto es necesario considerar que, además del lugar del impacto de la mirada, la posición de la cabeza dentro del entorno es dinámica, por lo que hay que tenerla en cuenta para reconocer las fijaciones. Dado que el algoritmo depende de una serie de parámetros, se ha realizado una metodología experimental, fundamentada en la optimización conjunta de un set de variables, para encontrar los valores óptimos de estos parámetros. El algoritmo ha sido publicado para ser utilizado por la comunidad científica.

En segundo lugar, se estudia la señal galvánica de la piel, que mide

la activación del sistema simpático. Esta señal suele presentar artefactos de movimiento que distorsionan la información que se pueda extraer de la misma. Debido a la dificultad que supone la detección y corrección de artefactos de forma automática, esta tarea se ha realizado mayoritariamente de forma manual. En esta tesis se propone el primer algoritmo de detección y corrección automática de estos artefactos, basado en modelos de aprendizaje profundo. Se desarrolla un modelo basado en redes convolucionales con una capa de entrada de *long short-term memory* para la detección de artefactos, mejorando los resultados del estado del arte. Para la corrección de artefactos se utiliza un algoritmo de regresión que reemplazaría la señal afectada. Los resultados de este trabajo muestran que la descomposición física de la señal corregida manualmente y la descomposición física que podría realizar un experto, no guardan diferencias significativas. Por el contrario, ambas sí la tienen frente a la señal original con artefactos. El modelo se ha hecho público para su uso en sistemas con procesamiento automático.

A continuación, se desarrolla una experimentación para la evocación y detección de emociones. Esta experimentación está basada en una conversación a tiempo real y por voz con un humano virtual. Este ha sido desarrollado utilizando la tecnología más avanzada de inteligencia artificial, como son los modelos generativos de lenguaje, transcripción de voz, síntesis de voz y sincronización labial. La experimentación se ha realizado en un entorno de realidad virtual semi-inmersiva en el que el humano virtual se muestra a través de una proyección a tamaño natural. El prototipo ha sido validado técnicamente y se han analizado las dinámicas comunicativas entre el humano virtual y el sujeto. Además, se ha evaluado la naturalidad y el realismo de las conversaciones, así como las emociones elicidadas en los sujetos.

Por último, se ha utilizado el prototipo para el reconocimiento de síntomas depresivos utilizando bio-marcadores de *eye-tracking* y la actividad electrodermal. Para ello, se ha realizado un experimento con 98 sujetos y desarrollado una metodología de análisis y validación basada en aprendizaje automático para realizar predicciones. Los modelos han alcanzado una precisión de 0.733 y una tasa de reconocimiento de pacientes no depresivos de 0.828. Además, se ha explorado el reconocimiento de naturalidad y realismo de las conversaciones, así como las emociones evocadas.

El trabajo desarrollado en esta tesis presenta contribuciones relevantes,

tanto para el campo de la computación afectiva, como para otras áreas afines como el procesamiento de señales y las interacciones humano-máquina. Las herramientas desarrolladas junto con modelos novedosos de la inteligencia artificial consiguen realizar una experimentación, en entornos de dinámicas sociales humanas, nunca antes diseñada. Los resultados muestran como este trabajo es capaz de modelizar información tan relevante como emociones y síntomas depresivos de la persona con la que se habla. Campos como la psicología, medicina o educación pueden utilizar muchas de las herramientas desarrolladas en esta tesis con tal de aportar más información en la toma de decisiones o en la interacción social.