

Captura multimodal para la inferencia de estados emocionales aplicados a contextos de computación afectiva

Jorge Ierache , Iris Sattolo , Gabriela Chapperon, Rocío Ierache, Fernando Elkfury, Carlos Barrionuevo , Facundo Nervo

Instituto de Sistemas Inteligentes y Enseñanza Experimental de la Robótica (ISIER)

Escuela superior de Ingeniería, Informática y Ciencias Agroalimentaria

Secretaría de Ciencia y Tecnología Universidad de Morón Cabildo 134, Buenos Aires, Argentina

{jierache,isattolo,gchapperon}@unimoron.edu.ar

Resumen

En este trabajo se enuncian el contexto de desarrollo en materia de I+D aplicada a la captura multimodal de emociones en diversos ámbitos de aplicación (virtuales, simuladores, robots, gastronomía), los proyectos vinculantes y los resultados publicados, como así también las líneas de I+D+I, la formación de RRHH. Finalmente, se expresa el reconocimiento al Ing. Hugo Rene Padovani, quien apoyara la formación de docentes investigadores del área informática en la UM .

Palabras clave: Computación Afectiva, Estado emocional, captura multimodal.

Contexto

Esta investigación aplicada se desarrolla en el contexto del Proyecto de Investigación Científica Tecnológica Orientado (PICTO) aprobado por la Agencia Nacional de promoción de la investigación, el desarrollo tecnológico y la innovación (ANPCyT), denominado “Influencias del estado biométrico emocional de personas interactuando en contextos de entornos simulados, reales e interactivos con robots” PICTO UM 2019-00005. El mismo se desarrolla dentro del Instituto de Sistemas Inteligentes y Enseñanza

Experimental de la Robótica de la Universidad de Morón (ISIER-UM) y está auspiciado por la Secretaría de Ciencia y Tecnología con una duración de tres años. Este trabajo se sustenta sobre las bases iniciales de las investigaciones realizadas en el marco de los proyectos: “Influencias del estado biométrico-emocional de personas interactuando en contextos de entornos virtuales” Ping/17-03-JI-004,(2017), el proyecto denominado “Explotación de datos EEG y parámetros fisiológicos de usuarios interactuando en contextos virtuales” DC-019 80020190100007 UM (2019) , el proyecto “Valoración Emocional Multimodal aplicada en contextos gastronómicos” PIO 80020190300016 UM (2020).

Introducción

El estudio de los estados afectivos en las personas es un campo que se aplica en diferentes realidades tales como la medicina, educación, conducción de automóviles y otros, con el fin de solucionar ciertos problemas aparejados por el dominio. A lo largo de la historia, la intuición y los trabajos de muchos profesionales de diversas disciplinas han indicado la importancia de las emociones, así como la necesidad de conocerlas y aprender a relacionarnos con éstas de la forma más coherente posible. Como ejemplo se puede citar a Darwing en 1872, James y Lang en 1884 y otros pensadores, quienes ya en su época

argumentan sus teorías sobre cómo se producen las emociones. En el campo de la computación se habla de proveer a las máquinas de emociones, originando un nuevo campo denominado computación afectiva. S. Baldasarri [1] plantea que los sistemas “afectivos”, deben ser capaces de capturar y reconocer los estados emocionales del usuario a través de mediciones sobre señales generadas en la cara, la voz, el cuerpo, o cualquier otro reflejo del proceso emocional que se esté llevando a cabo. Albert Mehrabian en sus investigaciones afirma que en una comunicación la palabra hablada solo contiene el 7% del significado, mientras que el tono de voz un 38% y las expresiones faciales nada menos que el restante 55% [2]. Existen distintos enfoques para representar las emociones, el enfoque dimensional [3] y el enfoque categórico. El enfoque categórico, desarrollado inicialmente por el psicólogo Paul Ekman [4], afirma que existe un conjunto de seis emociones básicas y universales que no están determinadas por las culturas. Luego actualizó este conjunto a siete emociones, sumando el desprecio. Estas son: a) enojo, b) asco, c) felicidad, d) miedo, e) tristeza, f) sorpresa, g) desprecio. Por otro lado, el enfoque dimensional establece que los estados afectivos se distribuyen en un espacio continuo, cuyos ejes dimensionales indican una cuantificación de una característica. A pesar de que actualmente no existe un consenso acerca de lo que es la definición de emoción, si existe un acuerdo en que las emociones pueden clasificarse en tres dimensiones continuas: 1) Valencia (Valoración): un constructo bipolar, que va de agradable a desagradable, 2) Excitación (Arousal): que va de calmado hasta activado (o excitado), y 3) Dominancia (Control). Los métodos utilizados en la computación afectiva se basan en el procesamiento de los datos recopilados y etiquetados obtenidos mediante la ejecución de experiencias en laboratorios, donde se pueden

obtener etiquetas (definidas con categorías predefinidas o de forma dimensional) de diferentes maneras: durante la interacción con el sistema o retrospectivamente, o por el investigador durante o después de la interacción. De acuerdo con los tipos de datos que se utilizan en los sistemas propuestos pueden clasificarse en sistemas unimodales, los que exploran una sola fuente de datos, y multimodales, los que combinan dos o más fuentes de datos. Como trabajos unimodales se puede citar a: en [5] se captura el rostro a través de videos y en [6],[7] captura de rostro con imágenes. Como trabajos multimodales se puede citar a: en [8] agregan a la captura del rostro la posición de la cabeza, en [9] utilizan cuestionarios, teclado y micrófono, en [10] captura de teclado y Electroencefalograma (EEG). Para deducir el estado emocional de un individuo en un contexto de captura multimodal (de ahora en más MM) se tiene que registrar, simultáneamente, diversa información biométrica. En este contexto se está trabajando la construcción de un Framework MM en el cual se combinen distintas características, distintos sensores y distintas estrategias para una mayor fidelidad de los datos capturados. En trabajos previos del ISIER-UM se trabajó en: [11], [12], [13], [14] control de Robots, en [15] el control de artefactos aplicados a la domótica, y en [16] la lectura emocional del usuario, enfocando la lectura de la excitación y meditación. Se experimentó con BCI (Brain Computer Interface), en particular con EMOTIV [17], integrando la respuesta de biopotenciales eléctricos de individuos para el control de robots a través de Electromiograma (EMG), EEG y Electrooculograma (EOG), las cuales son bioseñales eléctricas generadas por los patrones de actividad de los músculos, el cerebro y los ojos del usuario respectivamente.

Líneas de Investigación, Desarrollo e Innovación

La hipótesis que guía este proyecto es: La

captura de estados emocionales de personas interactuando en distintos contextos (reales, simulados, virtuales o interactivos con robots) a través de un Framework MM, independiente del dominio, con la capacidad de registrar datos tanto fisiológicos como biométricos, contribuye a reflejar un estado emocional de mayor fidelidad que el obtenido cuando se utilizan pocos medios de recolección de datos. Para demostrar dicha hipótesis, se encuentran abiertas y en desarrollo las siguientes líneas de trabajo:

a) evolución del Framework MM: registro de datos biométricos y fisiológicos; integración distintas interfaces BCI (Neurosky y Emotiv), distintos sensores fisiológicos (presión arterial, pulsaciones, conductancia de la piel), y distintas herramientas que interactúen con estas interfaces.

b) inferencia emocional a través de la captura del rostro del usuario, correlacionándolo con su estado emocional empleando diversas APIS, por ejemplo, entre otras, Affective de MS, a fin de comparar con desarrollos reciente de autor llevados a cabo en el ISIER aplicando Machine Learning (ML) sobre las bases de técnicas de regresión logística.

c) Desarrollo de escenarios de aplicación (virtual, simuladores, gastronómico, interacción con robots). Para la validación emocional se aplica la prueba de Russel (modelo excitación-valencia), encuestas SAM (por sus siglas en inglés Self-Assessment-Manikin) propuesta por Bradley y Lang [18], banco de imágenes IAPS [19] (genéricas para distintos cuadrantes del circunflejo de Russel), además de aplicar las imágenes de OLAF [20] (alimentos y comidas) para el dominio gastronómico. En la actuación del usuario en el contexto virtual, se aplican casos de experimentación empleando el conjunto de datos recopilado por el Center for the Study of Emotion & Attention de la University of Florida en Gaines Ville, FL,

Estados Unidos. El conjunto contiene una base de datos de imágenes (IAPS) [19], que es utilizada en diversos estudios científicos que analizan emociones y es considerada un estándar de facto.

d) visualización y explotación. Se desarrollan consultas y filtros con integración de herramientas de explotación y descubrimiento sobre la base de datos que articula el Framework.

Objetivos y Resultados Esperados

Los objetivos planteados en el marco de este PICTO son: a) Integración de distintas herramientas BCI al Framework MM, b) Integración de distintos sensores fisiológicos (pulsera, conductancia de la piel), c) Captura y registración de ritmo cardiaco, d) Integración de Rostro a través de regresión logística, e) Integración de Voz, f) Integración de Test de Rusell y Encuestas SAM, g) Capturas ambientes (simulados o reales), h) Integración de herramientas para descubrir patrones, i) implantación de un espacio físico de inmersión emocional.

Resultados obtenidos

Hasta el momento los logros obtenidos se reflejan en el Framework MM, donde para su desarrollo se consideraron como entradas principales la captura de las ondas cerebrales, con el empleo del BCI (Neurosky, Emotiv), sus algoritmos, y la toma de “capturas de pantalla”, tanto del rostro del usuario interactuando con el contexto como las imágenes que observa el usuario. Se integra con Encuestas SAM. Se realizaron pruebas con ambientes estáticos de imágenes y ambientes dinámicos, juegos de computadora. Los resultados de las pruebas mencionadas fueron demostrados y publicados CACIC 2018 [21], CACIC 2019 [22] y CACIC 2020 [23], CACIC 2020 [24], y resultados en ambientes dinámicos en Revista Risti [25].

Recientes trabajos no publicados a la fecha obtuvieron resultados satisfactorios en el desarrollo de ML aplicado a la detección de emociones por voz; en particular este trabajo se orienta al diseño y desarrollo de redes neuronales para la clasificación de emociones en el discurso hablado, con la experimentación de diferentes métodos para convertir de un enfoque categórico de clasificación de emociones a uno dimensional.

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por tres investigadores formados, tres investigadores en formación, y dos alumnos investigadores. Se defendieron en esta línea de investigación una tesis de doctorado en colaboración con FIUBA-UNLP, dos tesis de grado. Actualmente se desarrollan cuatro tesis de grado en el área de computación afectiva transfiriendo los resultados de investigación a la cátedra de trabajo de diploma de la carrera informática.

Reconocimiento

Los docentes investigadores integrantes del ISIER quieren expresar su reconocimiento al **Ing. Hugo Rene Padovani** quien en vida contribuyó a la creación del ISIER [26], como así también a la formación de RRHH en investigación en el área de robótica autónoma, y en el campo de la computación afectiva; sirva este espacio para expresar nuestro profundo agradecimiento, a nuestro querido Hugo, quien siempre estará presente en nuestros corazones.

Referencias

1. Baldassarri, S. (15 de 9 de 2016). Computación afectiva: tecnología y emociones para mejorar la experiencia del usuario. Obtenido de <http://hdl.handle.net/10915/53441>
2. A. Mehrabian, "Communication without Words", *Psychology Today*, vol. 2, no 4, (68)
3. Posner, J., Russell, J., & Peterson, B. (2005). "The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology.", *Development and psychopathology*, vol. 17, n° 3, págs. 715-34, 2005, ISSN: 0954-5794. *Development and psychopathology*, 715-34 Vol 17 n° 3.
4. Ekman, P. (2005 cap 3). "Basic Emotions": En e. T.Dalgleish y M.J.Power, *Handbook of Cognition and Emotion*, (págs. 45-60). Chichester, UK: John Wiley & Sons, Ltd.
5. van der Haar, D. T. (2019). Student Emotion Recognition. *International Conference on Human-Computer interaction* (págs. 301-311). Springer Chan.
6. Zatarain Cabada, R., Barron Estrada, M. L., Halor-Hernandez, G., & Reyes-García, C. A. (2014). Emotion Recognition in Intelligent Tutoring Systems. *Mexican International Conference on Artificial Intelligence* (págs. 494-504). Mexico: Springer.
7. Wei-Long, Z., & Bao-Liang, L. (2015). Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks. En I. X. Library, *IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT*, VOL. 7, NO. 3, SEPTEMBER 2015 (págs. 162-175). IEEE.
8. Xu, R., Chen, J., Han, J., Tan, L., & Xu, L. (2019). Towards emotion-sensitive learning cognitive state analysis of big data in education deep learning-based. En *Computing* (págs. 1-16). Austria: Springer Viena.
9. Alepis, E., & Virvou, M. (2006). User Modelling: An Empirical Study for Affect Perception Through Keyboard and Speech in a Bi-modal User Interface*. *International Conference on Adaptive Hypermedia and adaptive Web-Based Systems*. Berlin, Heidelberg: Springer.

10. Calot E., Ierache J., y Hasperu W. Robustness of keystroke dynamics identification algorithms against brain-wave variations associated with emotional variations. En *Intelligent Systems and Applications*, vol. 1037 de *Advances in Intelligent Systems and Computing*. Springer, 2019. ISBN 978-3-030-29515-8. doi: 10.1007/978-3-030-29516-5 15
11. Ierache, J., Dittler M., Pereira G., García Martínez R, "Robot Control on the basis of Bio-electrical signals" XV CACIC 2009, UNJu, ISBN 978-897-24068-3-9
12. Ierache J, Pereira G, Iribarren J, Sattolo I, "Robot Control on the Basis of Bio-electrical Signals": "International Conference on Robot Intelligence Technology and Applications" (RiTA 2012). Series *Advances in Intelligent and Soft Computing* of Springer.
13. Ierache., J, Pereira., G, Sattolo., I, Guerrero., A, D'Altto J, Iribarren, J. Control vía Internet de un Robot ubicado en un sitio remoto aplicando una Interfase Cerebro-Máquina". XVII CACIC 2011, UNLP, ISBN 978-950-34-0756-1, p 1373-1382.
14. Ierache J, Pereira G, Iribarren J "Demostración de los resultados en la integración de Interfases Lectoras de Bioseñales aplicadas al Control de un Robot" VII TEYET 2012 UNNOBA, 2012, demos educativas. ISBN 978-987-28186-3-0.
15. Ierache., J, Pereira., G, J, Iribarren Framework for Brain Computer Interface implemented to control devices in the context of home automation XIX CACIC 2013, CAECE Mar del Plata, ISBN 978-897-23963-1
16. Ierache J, Nervo F, Pereira G, Iribarren J Estado emocional centrado en estímulos, aplicando Interfase cerebro-maquina, XX CACIC (Buenos Aires, 2014) ISBN 978-987-3806-05-6
17. Emotiv Epoc Vigente 2018 <http://www.emotiv.com/>.
18. Bradley, M. M. & Lang, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59 (1994).
19. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. International Affective Picture System (IAPS): Technical Manual and Affective Ratings. NIMH Cent. Study Emot. Atten. 39–58, <https://doi.org/10.1027/0269-8803/a000147> (1997).
20. Miccoli, L., Delgado, R., Rodríguez-Ruiz, S., Guerra, P., García, E., & Fernández-Santaella, M. C. OLAF, the Open Library of Affective Foods in ADULTS:
- 21 J. Ierache, R. Nicolosi, G. Ponce, C. Cervino, E. Eszter. Registro emocional de personas interactuando en contextos de entornos virtuales. CACIC 2018, La Plata, pp 877-886, ISBN: 978-950-658-472-6
22. Ierache, J., Ponce, G., Nicolosi, R., Sattolo, I., & Chapperón, G. (2019). Valoración del grado de atención en contextos áulicos con el empleo de interfase cerebro-computadora. CACIC 2019. Rio Cuarto, pp: 417-426, ISBN: 978-987-688-377-1
23. Ierache J. Nervo F, Sattolo I, Ierache R, Chapperón G. Propuesta de un Modelo Multimodal de valoración emocional en el marco de la computación afectiva aplicado en ambientes gastronómicos. CACIC 2020 La Matanza. En prensa
- 24 Barrionuevo Carlos, Ierache. J. Sattolo I "Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística" CACIC 2020 La Matanza. En prensa
25. Ierache. J. Sattolo I, Chapperón G. Framework multimodal emocional en el contexto de ambientes dinámicos. DOI. 10.17013/risti 40. 45-59 ISSN: 1646-9895
26. www.isierum.cl.biz vigente febrero de 2021