

Avances en Línea de Investigación Doctoral: Integración Escalable de Realidad Aumentada Basada en Images y Rostros

Nahuel Mangiarua¹[0000-0003-2674-7324], Jorge Ierache¹[0000-0002-1772-9186], Martin Becerra¹, Maria Jose Abasolo²[0000-0003-4441-3264]

¹ Universidad Nacional de La Matanza, La Matanza, Buenos Aires, Argentina
nmangiarua@unlam.edu.ar

² Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina

Resumen

El presente trabajo resume una de las líneas de investigación doctoral en curso radicada en el Grupo de Realidad Aumentada Aplicada la cual se orienta al desarrollo de una arquitectura y prototipo demostrador que integre de forma escalable la realidad aumentada basada en imágenes y la aumentación de rostros humanos.

Palabras clave: Realidad Aumentada, Reconocimiento de Rostros

Contexto

La investigación es desarrollada en el grupo de investigación de Realidad Aumentada Aplicada del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de La Matanza, en el marco de un proyecto de tesis doctoral.

Introducción

La realidad aumentada mediante teléfonos móviles es una de las subáreas de investigación de realidad aumentada que más está creciendo en la actualidad. Se define a un sistema de Realidad Aumentada (RA) como un sistema que permite a usuarios interactuar en tiempo real con contenidos virtuales asociados a objetos del entorno real [1],[2]. Podemos ampliar esta definición como un sistema que aumenta los sentidos de la persona (Visión, audio y tacto) para permitir la interacción con contenidos virtuales del mundo digital que son invisibles en el entorno real [3]. El grupo de realidad aumentada aplicada realizó diversos trabajos en el campo de realidad aumentada que permite a usuarios sin mayores conocimientos técnicos de RA poder crear catálogos virtuales aumentados agrupando contenidos virtuales asociados a marcadores de RA

como imágenes impresas[4],[5]. En el área de educación el grupo de investigación ha realizado diversas aplicaciones experimentales [6],[7] orientadas a juegos, como así también aplicaciones experimentales destinadas a la explotación de materiales didácticos para el área educativa [8],[9].

Se presenta en las secciones siguientes la línea de investigación de tesis doctoral, Integración Escalable de Realidad Aumentada Basada en Imágenes y Rostros.

Líneas de Investigación, Desarrollo e Innovación

Integración de RA escalable

La línea de investigación doctoral se orienta al desarrollo de una arquitectura y prototipo demostrador que integre de forma escalable la realidad aumentada basada en imágenes y la aumentación de rostros humanos.

En cuanto a la arquitectura, se define un pipeline basal para la aumentación de imágenes y otro para la aumentación de rostros humanos como se muestra en las figuras 1 y 2.



Figura 1: Pipeline para AR basada en imágenes



Figura 2: Pipeline para reconocimiento y aumentación de rostros

Para la aumentación de imágenes, primero se debe, en forma offline, buscar puntos de interés (POI) y computar descriptores robustos sobre los mismos para cada imagen que se quiera aumentar con contenido virtual. Luego, durante la explotación, se buscan los mismos POI sobre las imágenes que van siendo capturadas desde un flujo de video y se computan los mismos descriptores robustos para los cuales se busca correspondencias contra los cálculos previamente offline para las imágenes a aumentar. Si tenemos coincidencias suficientes, podemos efectivamente detectar la presencia en el video de una de las imágenes a aumentar y luego calcular su posición y rotación relativas a la cámara.

En el caso de la aumentación de rostros, análogamente se computan offline descriptores robustos para cada persona cuyo rostro quiera ser reconocido mediante el uso de una red neuronal convolucional (CNN) [10],[11],[12]. Luego se utiliza una otra CNN entrenada para detectar la posición de rostros humanos sobre imágenes que vengan de un flujo de video y para cada detección, se computa un descriptor con la primera CNN. Los descriptores encontrados en las

imágenes provenientes del flujo de video deben entonces ser buscados por coincidencias entre los calculados previamente de forma offline y así identificar al individuo.

Se observa que ambos pipelines comparten un paso en común, la búsqueda de correspondencias entre descriptores robustos. A su vez, si analizamos la complejidad computacional de ambos pipelines, encontramos que mientras la mayoría de los pasos dependen del tamaño de las imágenes a procesar el cual puede ser fijado a un nivel adecuado en cada implementación concreta. Sin embargo, el paso de búsqueda de coincidencias depende del número de imágenes o rostros a aumentar, volviéndose el cuello de botella al momento de escalar el sistema.

Resultados y Objetivos

Bajo estas consideraciones se propone entonces la utilización de algoritmos de búsqueda de coincidencias aproximadas (approximate nearest neighbour, ANN) que ofrezcan una complejidad computacional sub lineal y estabilidad ante el incremento en el volumen de datos a procesar. Existen en la literatura infinidad de algoritmos de ANN y es actualmente un campo de investigación activo.

Se toma como objetivo analizar una muestra significativa de dichos algoritmos para elegir el idóneo para esta tarea particular en el contexto de la RA. A su vez resulta necesario definir y

adaptar conjuntos de datos de prueba representativos del contexto de la RA, seleccionando o creando las métricas específicas necesarias.

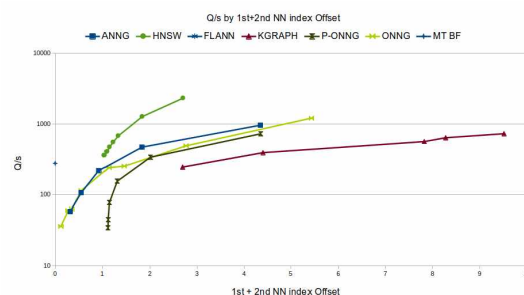


Figura 3: Análisis de la curva precisión sobre velocidad utilizando métricas y juegos de datos propuestos.

Como resultado preliminar observamos en la figura 3 el trade off entre la precisión y la velocidad de distintos algoritmos de ANN. Las mediciones son generadas utilizando una adaptación del set de datos SIFT1M[13] para adecuarlo a las necesidades de la RA mientras que la precisión se mide como el error promedio que comete cada algoritmo al encontrar la primera y la segunda mejor coincidencias para cada elemento.

Referencias

1. Papagiannakis G., Gurminder S. y Nadia M. T., "A survey of mobile and wireless technologies for augmented reality systems." *Comput. Animat. Virtual Worlds*, vol. 19, no. 1, pp. 3-22, 2008.
2. Azuma R. T., *The Most Important Challenge Facing Augmented Reality*, *Presence Teleoperators Virtual Environ.*, vol. 25, n.o 3, pp. 234-238, dic. 2016.

3. Specht M., Ternier S. y Greller W., “Dimensions of mobile augmented reality for learning: a first inventory,” *Journal of the Research for Educational Technology (RCET)*, vol. 7, no. 1, pp. 117-127, 2011
4. Ierache J., Mangiarua N., Verdicchio N., Becerra M., Duarte N., Igarza S. “Sistema de Catálogo para la Asistencia a la Creación, Publicación, Gestión y Explotación de Contenidos Multimedia y Aplicaciones de Realidad Aumentada”. CACIC 2014 Red UNCI ISBN 978-987-3806-05. 2014
5. Ierache J., Mangiarua N., Bevacqua S., Verdicchio N., Becerra M., Sanz D., Sena M., Ortiz F., Duarte N., Igarza S. “Development of a Catalogs System for Augmented Reality Applications”. World Academy of Science, Engineering and Technology, International Science Index 97, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 9(1), 1 - 7. ISSN 1307:6892. 2015
6. Becerra M., Sanz D., Igarza I, Mangiarua N., Ierache J. “Sistema de Catálogo Virtual Aumentado. Integración de Framework Especializado orientado a Material Didáctico” TEYET, pp 350-356, ISBN 978-950-656-154-3. 2015
7. Mangiarua N., Ierache J., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Igarza S. “Herramienta de Realidad Aumentada para la explotación de material didáctico tradicional”. TE&ET. E-Book. ISBN 978-987-24611-1. 2014
8. Verdicchio N., Sanz D., Igarza S., Mangiarua N., Montalvo C., Ierache J. “Sistema de Catalogo Virtual Aumentado Integración de Framework Especializado Orientado a Juegos Didácticos”. TE&ET, pp 597-604, ISBN 978-987-3977-30-5. 2016
9. Ierache J., Mangiarua N., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Igarza S. “Herramienta de Realidad Aumentada para facilitar la enseñanza en contextos educativos mediante el uso de las TICs”. *Revista Latinoamericana de Ing de Software*, 1(1): -3, ISSN 2314-2642. 2014
10. Parkhi, O.M., Vedaldi, A., Zisserman, A.: Deep face recognition. In: *BMVC* (2015)
11. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embed-ding for face recognition and clustering. CoRRabs/1503.03832 (2015), <http://arxiv.org/abs/1503.03832>
12. Taigman, Y., Yang, M., Ranzato, M., Wolf, L.: Deepface: Closing the gap to human-level performance in face verification. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23-28, 2014.* pp. 1701–1708. IEEE Computer Society (2014). <https://doi.org/10.1109/CVPR.2014.220>, <https://doi.org/10.1109/CVPR.2014.220>
13. Jégou, H., Douze, M., Schmid, C.: Product quantization for nearest neighbor search (2011)