

Uso de marcadores fiduciarios para la obtención de la posición y orientación de objetos en la escena para un simulador de soldadura usando realidad aumentada

Emiliano Churruca¹, Hernán Guzmán¹,

Fernando Puricelli¹, Leonardo Mottillo², Agustín Pérez², Guillermo Zarate¹, Juan Miguel Santos¹

¹Centro de investigación y desarrollo en informática aplicada (CIDIA), ²Grupo de Metalurgia en Soldadura y Procesos Avanzados (GMSyPA), Universidad Nacional de Hurlingham

{emiliano.churruca@unahur.edu.ar, hernan.guzman@estudiantes.unahur.edu.ar,

fernando.puricelli@unahur.edu.ar, leonardo.mottillo@unahur.edu.ar, agustin.tomas.perez@outlook.com.ar,

guillermo.zarate@unahur.edu.ar, juan.santos@unahur.edu.ar}

RESUMEN

Este trabajo expresa la línea de investigación y desarrollo en el uso de marcadores fiduciarios para la obtención de la pose de objetos en el espacio para un simulador de soldadura basado en realidad aumentada. Dado que hay diferentes alternativas de sistemas de marcadores, se hace una revisión de los mismos y se argumenta la motivación de usar la librería Aruco. Posteriormente analizamos los distintos factores que intervienen en la definición de los parámetros de uso: tamaño de los marcadores, cantidad de bits usados para su codificación y cantidad de marcadores usados. A partir de esto, se implementa un programa para generar diccionarios de marcadores y se propone una serie de experiencias para la definición de los parámetros mencionados. Finalmente, algunos resultados preliminares son mostrados.

Palabras claves: Realidad Aumentada, marcadores fiduciarios, Simulador de soldadura

CONTEXTO

La línea de trabajo descripta en este artículo es parte del proyecto PIUNAHUR 80020220100006 “Desarrollo de un prototipo de simulador de soldadura para capacitación en la técnica de soldado, en Argentina, para instituciones de formación técnica, tecnicaturas e ingenierías que incluyan cursos o materias de capacitación en soldadura” financiado por la Universidad Nacional de

Hurlingham (UNaHur) e incorporado al Banco Nacional de Proyectos de Desarrollo Tecnológico y Social del MINISTERIO DE CIENCIA, TECNOLOGÍA E INNOVACIÓN por resolución RESOL-2023-1119-APN-SACT#MCT. Este proyecto es llevado a cabo en el Centro de investigación y desarrollo en informática aplicada (CIDIA- UNaHur) con la colaboración del Laboratorio de Ingeniería en Metalurgia- UNaHur.

1. INTRODUCCIÓN

Los simuladores que usan realidad aumentada (RA) tienen como ventaja sobre los que usan exclusivamente realidad virtual (RV) que permiten al usuario, tener una percepción completa de cualquier tipo de escenario, en contraste con la RV restringida a un conjunto acotado de situaciones previamente establecidas. La RA ha sido usada como recurso de apoyo para procedimientos de diseño, ensamblaje y manufactura como así también para la construcción de simuladores para la enseñanza de prácticas industriales, en particular para el aprendizaje de técnicas de soldado [1] que se conocen como simuladores de soldadura. Esto simuladores, hace casi dos décadas, están presentes en industrias de todo el mundo y fueron introducidos, a lo largo del tiempo, por fabricantes de diversas empresas como [2], [3] y [4], entre otras.

No hay un estándar en lo que respecta a las técnicas usadas para simular las condiciones reales que experimenta un soldador durante la soldadura. Fast K. y colaboradores [5] proponen un sistema donde

una torcha real está asociada a un dispositivo que sensa, en los tres ejes, la fuerza ejercida sobre ella y un sistema de seguimiento de 6 grados de libertad en la cabeza del soldador (casco) y en la torcha. También, en la misma línea de trabajo, Claude Choquet [6] propone ARC+. Este simulador dispone de un dispositivo que realiza un seguimiento en 6 ejes de la mano y la cabeza del soldador y, una máscara con un visor para que el usuario visualice la escena reconstruida en forma virtual. Hay numerosos trabajos que reportan las características de diferentes simuladores de soldadura y un estudio bastante completo de estos se puede encontrar en [1].

Un aspecto de especial interés en este tipo de simuladores es el que refiere a cómo hacer que el usuario tenga la sensación de estar inmerso en un ambiente de soldadura. El problema de la inmersión en realidad aumentada implica poder reconocer objetos de la escena 3D y obtener su posición y orientación en el espacio. Esto permite sobre imprimir objetos virtuales, medir distancias y ángulos. Estas operaciones deben realizarse en un ciclo de ejecución breve para evitar que el usuario pierda la sensación de inmersión. A tal fin, se pueden usar sensores inerciales [7], o estaciones de posicionamiento, como así también marcas en los objetos de interés para recuperar la pose de los mismos. El estado del arte reconoce a los marcadores fiduciaros como el tipo de marcas más adecuadas para este tipo de tareas y además representan una opción más económica comparada con los sensores o dispositivos mencionados. Los marcadores fiduciaros son patrones gráficos con códigos específicos que se utilizan como puntos de referencia para detectar y rastrear la posición y orientación de los objetos en el espacio. Son imágenes bidimensionales con diseños únicos y estructuras que pueden ser fácilmente identificadas por sistemas de visión por computadora.

El uso de este tipo de marcadores fue evolucionando a partir de la aparición de nuevos diseños gráficos y de la posibilidad de definir diccionarios de los mismos (un diccionario es el conjunto de marcadores a ser reconocidos en un sistema de RA). ArtoolKit

[8] es un método basado en marcadores que son figuras cuadradas con un recuadro negro y con una imagen a reconocer en su interior. Como una evolución de ArtoolKit, Fiala, M [9] propuso ARTag que usa marcadores, basados en figuras cuadradas de 10 x 10 cuadraditos, de los cuales se utilizan 2 de ancho para los bordes y 6 x 6 para una región central que codifica al marcador. Su diccionario cuenta con 2002 marcadores diferentes.

La codificación usando códigos binarios en el interior del marcador se ha extendido y un número de técnicas han surgido tales como Matrix [10], ArtoolKit Plus [11] y Aruco [12] entre otras. Estas técnicas se diferencian en: 1) cómo detectan los bits de codificación del marcador, 2) la distancia de Hamming entre códigos de diferentes marcadores, y 3) la capacidad de manejar la oclusión (cuando un objeto se interpone en el campo de visión de la cámara e impide visualizar el marcador completo). También, cada una de estas técnicas se diferencia en el tipo de diccionario que usa teniéndose en cuenta si el mismo ya está pre establecido o puede ser generado por el usuario, la cantidad de marcadores que posee y la distancia de Hamming entre sus marcadores.

Por los motivos recién expresados, un examen de las distintas posibilidades condujo a la elección de la técnica de Aruco para nuestro desarrollo de realidad aumentada por los siguientes motivos: 1) la librería que implementa el método de detección de marcadores permite especificar la cantidad de bits para la codificación y el tamaño físico del marcador, 2) está documentado el método de generación de diccionarios lo cual permite hacer una implementación propia y generar los diccionarios *ad-hoc* para la aplicación, y 3) es compatible con librerías como OpenCV y OpenGL, las cuales son algunas de las que se usan en nuestro desarrollo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La utilización de Aruco para la inclusión de RA en el desarrollo del simulador de soldadura implica tener en cuenta las características propias de problema. Los

objetos de interés para la simulación son la probeta (material a soldar) y la torcha (pistola de soldado). La distancia de ambos objetos varía entre 30 cm y 60 cm. De acuerdo al campo de visión de la cámara y su resolución, dependerá el tamaño del marcador (medido en cm). Por otro lado, cuanto más lejos esté un marcador mayor es la probabilidad de no identificarlo. Desde este punto de vista, conviene usar marcadores grandes y con pocos bits en su codificación (por ejemplo, un marcador de 4 cm de lado con 3 x 3 bits). Sin embargo, en la tarea habitual del soldador, el guante y la torcha pueden ocluir sensiblemente a la probeta y por lo tanto será conveniente que la probeta tenga la mayor cantidad de marcadores posible. Hay que tener en cuenta que, cuando se incrementa la cantidad de marcadores utilizados, es necesario aumentar la cantidad de bits de su codificación si se quiere preservar una distancia de Hamming adecuada entre ellos. Pero, cuanto mayor cantidad de marcadores se usen más pequeño debe ser el tamaño forzando a disminuir la región dentro del marcador ocupada por cada bit de la codificación. Este análisis refleja una relación de compromiso entre el tamaño del marcador, la cantidad de marcadores necesarios y la cantidad de bits usados para la codificación.

Hay, además, otro aspecto relacionado con la iluminación y la calidad de la cámara usada en cuanto a su relación señal/ruido. La identificación de las esquinas del marcador define no solo la ubicación del mismo en la escena (traslación) sino también su rotación respecto al centro de coordenadas de la cámara. El procesamiento de la imagen para esta tarea se ve afectada por el ruido. Por ejemplo, si preservamos la cámara y el marcador fijos, la identificación de los bordes del marcador puede variar de cuadro en cuadro del video. Por este motivo, es deseable que la posición y orientación de un objeto no sean calculadas a partir de uno o unos pocos marcadores sino como un promedio de un grupo de ellos, lo más numeroso posible. Este hecho es parte de la relación de compromiso detallada más arriba. Finalmente, la forma de determinar la rotación y traslación que tiene el

sistema de coordenadas del marcador respecto al sistema de coordenadas de la cámara es mediante un método de estimación de pose [13]. En particular, la librería Aruco usa la función `SolvePnp()` implementada en OpenCV [14]. Tanto esta técnica como otras similares sufren, bajo determinadas circunstancias, una ambigüedad cuya resolución se traduce, usualmente, como una inversión del eje z en el sistema de coordenadas del marcador [15].

Por lo tanto, la línea de investigación y desarrollo planteada consiste en tratar con cada una de las situaciones mencionadas más arriba y encontrar una solución, ya sea experimental o algorítmica. Como parte de la labor, también se debe implementar un software para la generación de diccionarios donde se pueda especificar la cantidad de bits para su codificación (nxn), la distancia de Hamming mínima admitida y la cantidad de marcadores requeridos para el diccionario.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se implementó el algoritmo propuesto por Garrido et al. [12] para generar 4 diccionarios de 20 marcadores cada uno, pero usando 4 diferentes cantidades de bits para su codificación. Para marcadores de 3x3, 4x4, 5x5 y 6x6, la mínima distancia de Hamming obtenida fue 2, 6, 10 y 15 respectivamente. A partir de esto, podemos descartar la posibilidad de usar marcadores de 3x3 puesto que el algoritmo de corrección de errores durante la identificación de cada marcador solo puede corregir 1 bit y conduce a confusión durante la identificación.

Para tratar con el ruido en la imagen, establecemos la posición de cuatro vértices en el espacio 3D de la escena para dibujar un rectángulo sobre la imagen 2D basado en la proyección de dichos vértices. Cada proyección se calcula a partir del vector de traslación y la matriz de rotación de un marcador. En el caso que el cálculo se realice a partir de más de un marcador, la proyección sobre la imagen 2D que se toma es el promedio obtenido a partir de cada marcador.

Para esto, se dispuso de 4 tableros, cada uno con una hoja A4 y los vértices establecidos coinciden con los vértices de dicha hoja. En cada tablero se usaron diferentes cantidades de marcadores de 2 cm de lado cada uno: 5, 10, 20 y 40. La codificación usada fue 5x5 bits. Se obtuvo el valor de la proyección de cada vértice para 500 imágenes (cuadros de video) sin alterar la posición de la cámara ni del tablero. Esta experiencia se realizó 5 veces para cada uno de los tableros y se calculó la desviación estándar y la desviación máxima (respecto al promedio) para cada una de los 4 vértices proyectados en la imagen 2D. En la figura 1 vemos como disminuye la dispersión de la proyección calculada sobre un vértice de la hoja en función de la cantidad de marcadores usados en los tableros.

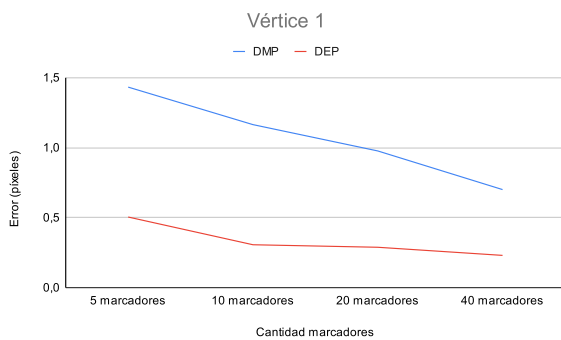


Figura 1. Desviación de la proyección de un punto en el espacio sobre la imagen 2D [en píxeles] vs. cantidad de marcadores en los tableros. El trazo azul corresponde a la distancia máxima respecto al promedio y el trazo rojo corresponde a la desviación estándar. El gráfico corresponde a uno de los vértices proyectados.

Con este resultado encontramos que para minimizar el error sobre la proyección en la imagen 2D es conveniente usar la mayor cantidad de marcadores posibles. Sin embargo, la cantidad de marcadores tiene una relación inversa con el tamaño en centímetros de cada marcador. Para determinar el tamaño mínimo del marcador en centímetros y cuál debe ser la codificación mínima en bits se llevó a cabo el siguiente experimento. Se creó un diccionario de 80 marcadores de 4x4 del cual se extrajo 40 al azar (sin discriminar la distancia de Hamming entre

ellos). Se armaron 2 tableros con los 40 marcadores, el primero con marcadores de 1 cm y el segundo de 2 cm. Se realizaron 4 experiencias cambiando el tablero y la distancia de la cámara al tablero (30 cm y 60 cm). En el caso del tablero con marcadores de 1 cm, independientemente de la distancia de la cámara al tablero, se verificó marcadores no detectados. Para el tablero con marcadores de 2 cm, se detectaban todos pero se confundía la identificación de alguno de ellos. Esto nos permite descartar los marcadores con una codificación de 4x4.

Repetimos la experiencia pero con un diccionario de 5x5. En este último caso se encontró que en el tablero con marcadores de 1 cm aun había marcadores sin detectar. Sin embargo, con el tablero con marcadores de 2 cm, sea cual fuere la distancia de la cámara al tablero, todos los marcadores fueron detectados e identificados correctamente tal como se puede observar en la figura 2.

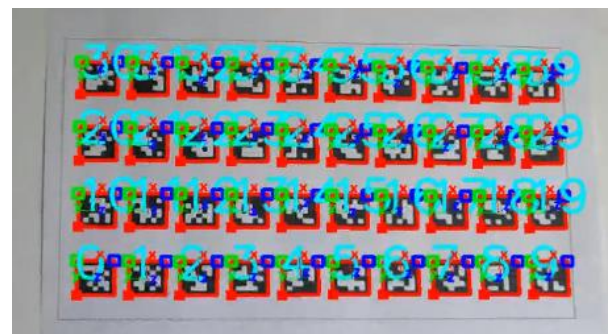
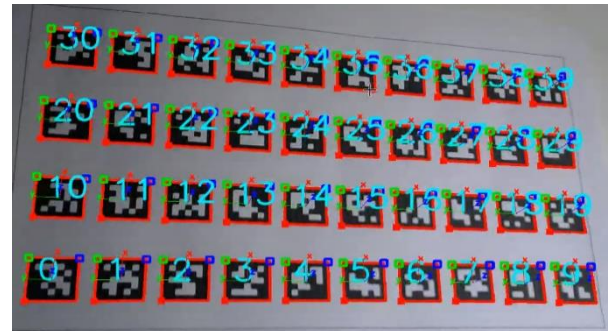


Figura 2. Dos capturas de pantalla mostrando la detección e identificación de todos los marcadores de 5x5, de 2 cm de lado en el tablero. Arriba a 30 cm de distancia de la cámara, abajo a 60 cm de distancia.

Actualmente se están llevando cabo nuevas experiencias y diseñando estrategias para

tratar con el resto de los problemas mencionados en la sección 2.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo del CIDIA se encuentra conformado por tres investigadores formados, dos investigadores estudiantes de doctorado, un estudiante de maestría y cuatro alumnos de grado que se encuentran realizando la tesis final de grado. Este trabajo constituye la idea más importante de la tesis para obtener el título de Licenciado en Informática, de los alumnos Emiliano Churruca y Hernán Guzmán.

5. BIBLIOGRAFÍA

- [1] Lavrentieva, O., Arkhypov, I., Kuchma, O., & Uchitel, A. (2020). Use of simulators together with virtual and augmented reality in the system of welders' vocational training: past, present, and future.
- [2] <http://weldsimulator.com/weldingsimulator>
- [3] <https://simbott.com/welding-simulator/>
- [4] <https://www.maquisol.com/categoria/nuevastecnologias/soldamatic/>
- [5] Fast, K., Gifford, T., & Yancey, R. (2004, November). Virtual training for welding. En *third IEEE and ACM international symposium on mixed and augmented reality* (pp. 298-299). IEEE.
- [6] Choquet, C. (2008). Arc+: Today's virtual reality solution for welders. *Internet Page, Jan, 1*(6).
- [7] Ni, D., Yew, A. W. W., Ong, S. K., & Nee, A. Y. C. (2017). Haptic and visual augmented reality interface for programming welding robots. *Advances in Manufacturing, 5*(3), 191-198.
- [8] H. Kato, M. Billingham, Marker tracking and HMD calibration for a video-based augmented reality conferencing system, en: *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, IWAR 099, IEEE Computer Society, Washington, DC, USA, 1999*, pp. 85–94.
- [9] Fiala, M. (2005, June). ARTag, a fiducial marker system using digital techniques. En *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 2, pp. 590-596). IEEE.
- [10] Rekimoto, J. (1998, July). Matrix: A realtime object identification and registration method for augmented reality. En *Proceedings. 3rd Asia Pacific Computer Human Interaction (Cat. No. 98EX110)* (pp. 63-68). IEEE.
- [11] D.Wagner,D.Schmalstieg, ARToolKit Plus for pose tracking on mobile devices, en: *Computer Vision Winter Workshop, 2007*, pp.139–146.
- [12] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Marín-Jiménez, M. J. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition, 47*(6), 2280-2292.
- [13] Eric Marchand, Hideaki Uchiyama, and Fabien Spindler. Pose Estimation for Augmented Reality: A Hands-On Survey. *IEEE Transactions on Visualization and Computer Graphics, 22*(12):2633 – 2651, December 2016.
- [14] https://docs.opencv.org/4.x/d5/d1f/calib3d_solvePnP.html
- [15] Eivazi Adli, S., Shoaran, M., & Sayyed Noorani, S. M. (2020). GSP n P: simple and geometric solution for P n P problem. *The Visual Computer, 36*, 1549-1557.