

Methods for Motion Detection in Paper Mechanisms for Tangible Augmented Reality Interfaces

Imelda Latapie, Laura Méndez Segundo, David Araujo Díaz, Miriam Pescador-Rojas

Published: 30 November 2024

Abstract

Augmented reality connects the physical world with digital environments, offering users enriched experiences for various applications, including entertainment, education, and marketing. There are proposals for tangible augmented reality interfaces, with paper mechanisms standing out due to their wide range of designs and low construction cost. These mechanisms simulate circular, lever, or linear movements along one or multiple axes, serving as tools for manipulating objects in virtual environments.

This article presents prototypes of tangible interfaces based on paper mechanisms utilizing sliding tabs and incorporating elements such as ArUco markers for detecting displacement along a single axis. The tests were conducted in a controlled environment with specific lighting conditions, progressively varying the distance between the camera and the proposed mechanisms. The results highlight the advantages and disadvantages of the employed techniques, demonstrating that good motion detection accuracy can be maintained. This facilitates the direct manipulation of virtual objects for augmented reality applications.

Keywords:

Tangible Interfaces, Augmented Reality, Paper Mechanisms, Markers, Computer Vision.

1 Introducción

En años recientes la realidad aumentada (RA) ha incrementado su presencia en diferentes ámbitos, como la educación, el entretenimiento, las publicaciones editoriales, la industria, la capacitación, el control de aplicaciones robóticas, o la creación artística. Se prevé que la RA aumentará su relevancia en un futuro cercano, en aplicaciones de telemedicina, el sector de la construcción, el soporte remoto, el diseño digital, la mercadotecnia y la publicidad. De acuerdo con el equipo de *MarketSplash*, se

estima que para 2027, el mercado de la RA y la realidad virtual será de 2,593 millones de usuarios [1].

La RA integra representaciones virtuales y elementos generados artificialmente con la percepción del mundo físico [3], es decir, se agrega información virtual sobre la realidad física que se capta a través de una cámara. Esta combinación de imagen real e información digital se proyecta sobre una pantalla que puede estar en un teléfono inteligente, una tableta, en gafas inteligentes o pantallas gigantes. Una interfaz de usuario tangible busca integrar información digital a objetos y entornos cotidianos [4]. Lo anterior, permite que los objetos físicos se conviertan en partes activas de la interfaz, cuando el usuario manipula directamente el objeto para interactuar con la información digital.

La RA tangible se refiere a interfaces de RA que están basadas en los principios de las interfaces de usuario tangibles, y combinan el manejo intuitivo de los objetos físicos con las capacidades de visualización que ofrece la RA, al superponer imágenes virtuales en el mundo real [5].

Existen antecedentes relevantes de interfaces tangibles de papel, como los Touchets [6], que son figuras planas de papel recortadas por el usuario, las cuales se registran en una aplicación para asociarles un movimiento y una acción. Una cámara de profundidad hace el seguimiento visual del objeto, de la mano y el dedo del usuario, para detectar el momento en que se manipula el objeto de papel. En este trabajo se asume que la mano es el único elemento tridimensional sobre la superficie, y hace un seguimiento de las figuras por su color.

En 2020, C. Zheng y otros [7] presentan una propuesta que consiste en estructuras de papel y marcadores que se imprimen en papel común, se cortan y se doblan para formar botones, barras deslizantes o de menú que son manipulados por el usuario para formar o distorsionar un marcador legible por la cámara. Estos mecanismos se convierten en teclas de instrumentos de música, controles de videojuegos o reproductores de música, que a pesar de su fragilidad resultan efectivos.

En 2022, H.-N. Hsu [8] propone un prototipo de cubo dividido en dos mitades que permanece unido desde el centro. Tiene dos estados para operar; en uno, permite la rotación del objeto virtual en cada eje, al girar las caras del cubo, y en el otro, genera el escalamiento al desplazar hacia arriba o hacia abajo el cubo. Para cambiar de estado, las mitades del cubo se giran en direcciones opuestas. Se calcula la transformación del objeto virtual dependiendo del cambio de posición del cubo. Esta propuesta utiliza una plataforma de software para controlar sistemas ópticos

Latapie I., Méndez L., Araujo D., Pescador M.,
Escuela Superior de Cómputo, Instituto Politécnico Nacional
Ciudad de México, México.
Email: ilatapiev@alumno.ipn.mx, {lmendez, daraujo,
mpescadorr}@ipn.mx

de captura de movimientos (OptiTrack Motive), que puede rastrear objetos con seis grados de libertad en tiempo real. La posición del cubo se captura con múltiples cámaras sincronizadas, distribuidas alrededor de la habitación, lo que proporciona datos de seguimiento muy precisos.

Los trabajos descritos exploran nuevas formas de participar en procesos de comunicación y control de acciones, abriendo áreas de oportunidad para el desarrollo de mejoras en el proceso de interacción, con interfaces accesibles para un público más amplio. Puede apreciarse que convertir un mecanismo de papel en un marcador para RA enfrenta retos relacionados con su detección y su seguimiento a través de una cámara, así como el procesamiento en tiempo real para ejecutar acciones entre los objetos de la realidad virtual y el mundo real.

Los mecanismos de papel son componentes bidimensionales interactivos, por ejemplo, una solapa o lengüeta, que al ser accionados generan un efecto de movimiento o transformación física en la escena o imagen a la que está asociado [9]. Estos mecanismos pueden crear efectos visuales y movimientos de algún elemento de papel en su entorno físico, a través de cortes, pliegues y ensamblajes precisos. Existen diferentes tipos de mecanismos de papel como: los discos rotatorios con efectos de movimientos circulares; los de cambio de escena disolvente que integran capas superpuestas que cambian gradualmente una imagen a otra; los que tienen planos levantados que incluyen tridimensionalidad y profundidad; los de pivote que generan movimientos complejos de otras piezas; o los de movimientos deslizables con lengüetas que tiran o empujan otros elementos de papel [10].

Un mecanismo de papel como soporte de una interfaz tangible ofrece ventajas por su peso ligero, su bajo costo, la posibilidad de cambiar de forma, y por la firmeza que adquiere al plegarlo. Estas características permiten crear interfaces interactivas, atractivas y generar experiencias de usuario novedosas y envolventes, como las descritas en [6] y [7]. Además, este tipo de mecanismos pueden funcionar como marcador para interactuar físicamente con la RA de manera controlada, para integrar elementos de realidad aumentada con los objetos físicos de la interfaz tangible que gestionen el despliegue y manipulación del contenido digital en tiempo real, por lo que es crucial identificar y rastrear objetos como etiquetas y marcadores.

Un marcador se describe como una imagen, ícono u objeto bidimensional que la cámara pueda reconocer, y proporciona una clave visual donde posicionar el contenido digital. Existen varios tipos de marcadores de referencia o fiduciales, y los de forma de base cuadrada binaria como por ARToolKit, AprilTags, ARTag, o ArUco se utilizan frecuentemente en el campo de la RA [14], [15], [16]. Los marcadores ArUco están delimitados por un borde negro externo, y cada marcador tiene, en la región interna, un patrón binario único que lo identifica en un diccionario. Además, puede verse como una matriz de $n \times n$ (donde n es cualquier entero, generalmente entre 4 y 7), y mientras que el borde externo facilita una detección rápida, la codificación binaria interna permite la identificación y la aplicación de técnicas de corrección de errores. Estos marcadores son parte de la biblioteca ArUco [17] de código abierto en lenguaje C++ que proporcionan como puntos de referencia sus cuatro esquinas los cuales permiten obtener la posición de la cámara.

En este artículo se presenta el desarrollo de dos prototipos de interfaces tangibles en forma de mecanismos de papel como marcador para controlar la interacción del usuario con un elemento digital de RA. Se describen dos técnicas para el reconocimiento y seguimiento de movimiento en un eje mediante mecanismos de papel basados en lengüetas deslizables que integran marcadores

ArUco. El contenido de este artículo se distribuye de la siguiente manera: en la sección 2 se presentan las propuestas de mecanismos de papel para la detección del desplazamiento de uno de sus elementos. La sección 3 se centra en las pruebas realizadas y los resultados obtenidos. En la sección 4 se describe una discusión de resultados y finalmente se presentan las conclusiones y las direcciones futuras del trabajo en la sección 5.

2 Propuestas de detección mecanismos de papel

En esta sección se describen dos propuestas de prototipos de mecanismos de papel basados en una lengüeta deslizable de papel, los cuales combinan ciertos elementos para detectar el movimiento de desplazamiento y conseguir precisión al variar la distancia entre una cámara y el mecanismo (véase Figura 1 y 2). Los mecanismos de papel se construyeron mediante cartulina Opalina Bright de 176 gramos. Aunque pueden elaborarse con cualquier cartulina rígida que permita una manipulación continua de la lengüeta.

2.1 Descripción del prototipo 1

El primer mecanismo de papel propuesto tiene una lengüeta deslizable que abre o cierra gradualmente una ventana, mostrando un área verde interior, como se muestra en la Figura 1. Además, se incorporó un marcador ArUco como punto de referencia para calcular el área verde visible en el desplazamiento de la ventana del mecanismo de papel.

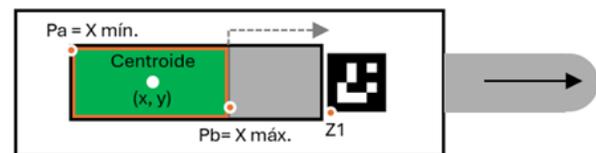


Figura 1. Primer prototipo de mecanismo de papel con máscara de color y un marcador ArUco como punto de referencia. La flecha punteada indica el movimiento que genera la lengüeta deslizable.

Para identificar el área en color verde se usaron las bibliotecas de OpenCV y Numpy de Python que ofrecen algoritmos optimizados para la detección y reconocimiento de objetos. Se capturó la imagen del mecanismo de papel y se convirtió al modelo de color HSV (Hue, Saturation, Value) para eliminar la presencia de piel (de manos o dedos) a través de una máscara con el umbral del color verde. El modelo HSV representa los colores de una manera más cercana a cómo los humanos percibimos el color. En esta propuesta, se separó la información de color (matiz y saturación) de la información de brillo (valor), lo que permitió una detección del color de piel más robusta ante cambios de iluminación, ya que los tonos de piel suelen ocupar un rango específico de matices y saturaciones. Adicionalmente, se hicieron algunos ajustes en el umbral de detección del color verde para descartar zonas muy pequeñas que no correspondan al mecanismo de papel.

Para el modelo de color HSV, usamos OpenCV con los siguientes valores. H: 0–180, S: 0–255, V: 0–255 [11].

En el umbral de detección del color verde, H1 (verde claro) y H2 (verde oscuro) representan los siguientes valores HSV seleccionados.

$$\begin{aligned} H1 &= (35, 50, 20), \\ H2 &= (80, 255, 255). \end{aligned}$$

Para determinar el centroide de la imagen se calcularon los momentos de la imagen con `cv2.moments()` y se utilizaron los momentos de orden 0 y 1 [2, 12]. M_{00} es el momento de orden cero, y representa el área del objeto. (Véase Figura 1)

$$\text{Centroide} = (x, y) = \left\{ \frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right\} \quad (1)$$

El cálculo del área resulta de utilidad para los mecanismos de papel que al incrementar el área mostrada no cambian el centroide, como es el caso de las ruletas o los de apertura de iris.

De manera adicional a la detección del área verde, se identificó el marcador y sus cuatro esquinas, y se tomó una de ellas como punto de referencia (Figura 1). Para la identificación del marcador, la imagen se transformó a escala de grises y se utilizó el diccionario predefinido `DICT_4X4_50` [13].

2.2 Descripción del prototipo 2

En el segundo prototipo, se eliminó la ventana en color, ya que se sustituyó por un método de detección y seguimiento basado en tres marcadores ArUco, véase la Figura 2. Los marcadores Pa y Pc son fijos y se utilizaron para determinar las coordenadas mínima y máxima en el eje X, representando el inicio y el final de la distancia total de desplazamiento. Mientras que el marcador Pb es móvil mediante la lengüeta deslizante y se asocia a un movimiento de desplazamiento en el eje X.

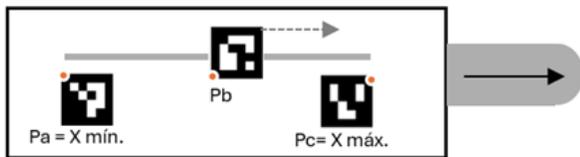


Figura 2. Segundo prototipo de mecanismo de papel que incluye dos marcadores ArUco fijos (Pa y Pc), y uno que es móvil (Pb) mediante la lengüeta deslizante, este movimiento se representa con la flecha punteada.

Cuando los marcadores ArUco Pa y Pb coinciden en el mismo punto es equivalente a no tener área visible en el primer prototipo. El caso contrario es cuando los marcados ArUco Pb y Pc coinciden lo que es igual a dejar visible toda el área en color verde del primer prototipo. Lo anterior permitió realizar una comparativa de precisión en movimiento para ambas propuestas.

2.3 Descripción del ambiente de pruebas

Las pruebas se realizaron utilizando la cámara web cuya resolución es de 1920 X 1080, vídeo estándar FHD de 2.07 MP, luminosidad 58.98% y 67617 colores, con el mecanismo de lengüeta y con luz artificial (como complemento a la iluminación natural) durante la mañana y la noche. Además, las imágenes capturadas por la cámara web se procesaron en una Raspberry Pi 4 modelo B, 2 GB RAM.

Las pruebas se realizaron en una habitación con iluminación mixta: luz ambiental proveniente de una ventana cerrada y cubierta con una cortina, y una fuente de luz artificial proveniente de una lámpara de mesa con un foco de luz cálida equivalente a 100 watts, colocada detrás de la cámara. En ambos casos se registraron resultados en diferentes momentos del día, por la mañana y la noche. Se utilizó una cámara web cuya resolución es de 1920 X 1080, vídeo estándar FHD de 2.07 MP, luminosidad 58.98% y 67,617 colores. Las imágenes capturadas por la cámara web se procesaron en una tarjeta Raspberry Pi 4 modelo B, 2 GB RAM. Se hicieron algunos ajustes en el umbral de detección del color verde para descartar zonas muy pequeñas que no corresponden al

mecanismo de papel. Los mecanismos de papel propuestos se fijaron a un bastidor de madera, a 30 centímetros sobre la superficie de una mesa. Mientras que la distancia de separación de la cámara fue 30, 50 y 70 centímetros. La distancia se midió con una cinta métrica, fijando el bastidor en cada posición.

La tabla 2 muestra los resultados que se obtuvieron de la estimación promedio en el desplazamiento para cada prototipo. Se realizaron 30 lecturas de la cámara en las distancias de separación definidas. Entre paréntesis se reporta la desviación estándar.

Área visible	Prot	Distancia		
		30	50	70
0%	1	14.0958 (7.5323)	16.0846 (8.1713)	12.4970 (7.8324)
	2	0.6479 (0.2976)	0 (0)	0.2268 (0.5495)
25%	1	21.3353 (16.2099)	17.1034 (8.7549)	12.1066 (7.3580)
	2	25.8499 (0.3017)	26.5881 (0.3322)	25.9533 (0.3933)
50%	1	34.1194 (28.8715)	27.2489 (16.7262)	20.5668 (20.6768)
	2	49.5538 (0.2866)	50.3410 (0.3081)	50.9639 (0.2678)
75%	1	63.4142 (42.7062)	42.5035 (39.6727)	18.4946 (23.2727)
	2	75.4575 (0.1891)	75.6404 (0.6323)	75.8512 (0.4101)
100%	1	49.6898 (41.9847)	40.0797 (40.4200)	30.2421 (36.7588)
	2	99.1780 (0.3080)	99.2665 (0.3266)	99.6254 (0.6142)

Tabla 1. Resultados de la estimación de desplazamiento en cada prototipo. Datos de salida del mecanismo a 30, 50 y 70 cm de distancia de la cámara y con 5 variaciones de desplazamiento. El valor entre paréntesis hace referencia a la desviación estándar al realizar 30 lecturas independientes.

3 Discusión de resultados

Las pruebas que se realizaron con los dos prototipos sirvieron para evaluar la precisión en el desplazamiento de los mecanismos, lo que deberá verse reflejado en la manipulación de un objeto virtual para una aplicación de RA. Con el primer prototipo, la máscara identificó correctamente el área verde visible del mecanismo de papel, manteniendo la precisión a pesar de cambios en la posición e inclinación del mecanismo. No obstante, el trazado del contorno y la localización de las coordenadas mínimas y máximas en el eje X dentro del área verde mostraron inconsistencias que dependían de la iluminación y la inclinación del mecanismo. Acercar o alejar el mecanismo de papel tendía a afectar los valores de las coordenadas, independientemente del desplazamiento de la ventana y del área verde visible. El marcador ArUco y las coordenadas de sus esquinas se identificaron con notable precisión. Sin embargo, las características de la cámara y las condiciones de iluminación capturaron múltiples zonas que fueron identificadas como verdes dentro del rango de la máscara, aunque en realidad eran de colores distintos (blanco, café o beige). Esto afectó el cálculo del desplazamiento de la lengüeta, ya que se identificaron múltiples zonas verdes con su centroide y sus coordenadas mínima y máxima de manera simultánea.

Al utilizar una máscara en el modelo de color HSV se reducen las operaciones de preprocesamiento. La iluminación juega un papel muy importante en la detección del objetivo, por lo tanto, en el posterior cálculo de los momentos y de las coordenadas en x mínima y máxima de área verde visible. Se observó que las dimensiones del área verde calculadas (y , por lo tanto, las coordenadas mínima y máxima de x , y las coordenadas del centroide en el área verde) varían no solamente por la posición de la lengüeta del mecanismo de papel, sino que también se ve afectada por la distancia de la cámara al dispositivo de papel.

Por otro lado, en los resultados del prototipo 2 se observan medidas muy precisas en todos los casos de prueba, a pesar de aumentar la distancia entre la cámara y los marcadores, lo que brinda un acoplamiento idóneo para el uso de gafas de realidad virtual o de RA.

4 Conclusiones y trabajo a futuro

En este artículo se presentaron dos métodos asociados a prototipos de mecanismos de papel para reconocer el movimiento de desplazamiento, en cada propuesta surgen diferentes ventajas. Por un lado, el uso de marcadores ArUco permite detectar con precisión los movimientos en el mecanismo de papel basado en lengüeta que se propone, ya que el marcador sirve como punto de referencia aun cuando cambian las condiciones en la escena que se visualiza, tales como la distancia entre el mecanismo y la cámara, la inclinación, o cambios en la iluminación. Además, el diccionario de marcadores nos permite manejar sin problema diferentes objetos virtuales que pueden interactuar en escenas de RA. Por otro lado, usar una máscara como el prototipo 1 brinda ventajas para la detección de áreas y sin importar la distancia e inclinación. Sin embargo, existen limitaciones cuando se varía la iluminación, ya que este afecta al tono de verde que detecta la cámara, de la misma manera las características de la cámara cambian los tonos e intensidad. Como trabajo a futuro se propone diseñar otros mecanismos de papel para detectar movimientos circulares o de palanca. Además, probar la interacción de los mecanismos al agregar objetos virtuales en una aplicación de RA y evaluar la experiencia de usuario en un producto final.

5 Referencias

- [1] MarketSplash, [En línea]. Available: <https://marketsplash.com/estadisticas-de-realidad-aumentada/>. [Último acceso: 30 5 2024].
- [2] E. Ramírez y Moreno Francisco, Algoritmos de Visión por Computador para un SBC, 2017.
- [3] A. Craig y W. Sherman, Understanding Virtual Reality. Interface, application, and design., Elsevier Science, 2003.
- [4] H. Ishii y B. Ullmer, «Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms,» Conference on Human Factors in Computing Systems-Proceedings, 12 9 1998.
- [5] M. Billingham, H. Kato y I. Poupyrev, «Tangible augmented reality,» ACM SIGGRAPH ASIA, 2008.
- [6] V. Becker, S. Kalbermatter, S. Mayer y G. Sörös, «Tailored Controls: Creating Personalized Tangible User Interfaces from Paper,» 2019.
- [7] C. Zheng, P. Gyory y E. Y.-L. Do, «Tangible Interfaces with Printed Paper Markers,» de 2020 ACM Designing Interactive Systems Conference, Eindhoven, 2020.
- [8] H.-N. Hsu, Paper-based Tangible User Interface as Interaction in Mixed Reality, 2022.
- [9] M. Serrano Sánchez, ¡Pop up! La arquitectura del libro móvil ilustrado infantil, Granada, 2018.
- [10] C. del Rosario Ruiz JR, Automated paper Pop-Up design: approximating shape and motion. Tesis doctoral., Singapore: Automated paper Pop-Up design: approximating shape and motion. Tesis doctoral., 2015.
- [11] A. Agrawal, «Color Swapping techniques in Image Processing,» 25 5 2021. [En línea]. Available: <https://towardsdatascience.com/color-swapping-techniques-in-image-processing-fe594b3ca31a>. [Último acceso: 31 5 2024].
- [12] ACTUMLOGOS, «Actumlogos, Visión por Computadora,» [En línea]. Available: <https://www.actumlogos.com/>. [Último acceso: 31 5 2024].
- [13] O. Kalachev. "Online ArUco marker sheet generator". Fodi @ GitHub. Accedido el 1 de julio de 2024. [En línea]. Disponible: <https://fodi.github.io/arucosheetgen/>
- [14] R. Şahin y M. Mercimek, «Fiducial Markers Aided Position Estimation for Vertical Landing », The European Journal of Research and Development, vol. 3, págs. 29-45, jun. de 2023, Accedido: 18-nov-2023.
- [15] A. González Martín, «Control de asistencia basado en reconocimiento robusto de fiduciales: aplicación para sistemas móviles.,» Accedido: 18-nov-2023, Licenciatura, Universidad de La Laguna, Escuela Superior de Ingeniería y Tecnología, jul. de 2015.
- [16] M. Košťák, A. Slabý. "Designing a Simple Fiducial Marker for Localization in Spatial Scenes Using Neural Networks," in Sensors, vol. 21, no. 16, 2021.
- [17] OpenCV. "OpenCV: Detection of ArUco Markers". OpenCV documentation index. Accedido el 26 de agosto de 2024. [En línea]. Disponible: https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html



© 2024 by the authors. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.