

CARACTERIZACIÓN DE UN SISTEMA OPERATIVO DE INTERFAZ DE USUARIO TRIDIMENSIONAL CONTROLADO POR OJOS, VOZ Y GESTOS

Features of a three-dimensional user interface operating system controlled by eyes, voice and gestures

John Fredy Robayo Ramírez

Universidad Libre de Colombia, Bogotá, Colombia,
johnf-robayor@unilibre.edu.co

RESUMEN

El artículo aborda la revolución tecnológica actual, y destaca la influencia de la inteligencia artificial en empresas líderes como Meta y Apple. Se centra en las innovadoras tecnologías lanzadas recientemente, como las Apple Vision Pro y Oculus Insight, que representan avances significativos en la realidad mixta y la realidad virtual. Detalla las especificaciones de las Apple Vision Pro, y las presenta como una “computadora espacial” que fusiona lo digital y lo físico, con características destacadas como la autenticación biométrica basada en iris y un sistema de seguimiento ocular y manual. Además, se discute la arquitectura del sistema y se resalta la importancia de enfoques multidisciplinarios para garantizar experiencias inmersivas.

Se explora también el funcionamiento de estas tecnologías, que incluyen la seguridad de autenticación de las aplicaciones, el seguimiento ocular y manual, y la capacidad de comprender la ubicación y tamaño de objetos del mundo real. Finalmente, se analizan diversos usos prácticos, desde aplicaciones médicas, como cirugías remotas y terapias, hasta su impacto en la industria y la educación, lo cual demuestra su potencial para generar experiencias inéditas y contribuir al bienestar general.

Palabras clave: inteligencia artificial, realidad mixta, realidad virtual, autenticación biométrica, seguimiento ocular, seguimiento manual, tecnologías inmersivas, arquitectura del sistema, aplicaciones médicas, industria, educación.

ABSTRACT

The article addresses the current technological revolution, highlighting the influence of artificial intelligence on leading companies such as Meta and Apple. It focuses on recently released innovative technologies, such as Apple Vision Pro and Oculus Insight, which represent significant advances in mixed reality and virtual reality. The article details the specifications of the Apple Vision Pro, presenting it as a “spatial computer” that fuses the digital and the physical, with notable features such as iris-based biometric authentication and an eye and hand tracking system.

Furthermore, the system architecture is discussed and the importance of multidisciplinary approaches to ensure immersive experiences is highlighted. The operation of these technologies is also explored, including application-level authentication security, eye and hand tracking, and the ability to understand the location and size of real-world objects. Finally, various practical uses are analyzed, from medical applications, such as remote surgeries and therapies, to its impact on industry and education, demonstrating its potential to generate unprecedented experiences and contribute to general well-being.

Keywords: Artificial intelligence, mixed reality, virtual reality, biometric authentication, eye tracking, hand tracking, immersive technologies, system architecture, medical applications, industry, education.

1. INTRODUCCIÓN

En la actualidad, los seres humanos viven en un mundo marcado por cambios tecnológicos casi diarios. Tras el notable crecimiento de la inteligencia artificial, esta ha establecido pautas que permiten la realización de lo que la mente humana puede imaginar y definir como objetivo. Un ejemplo destacado de este fenómeno se encuentra en algunos gigantes tecnológicos y pioneros en la creación de hardware y software, como son los casos de Meta y de Apple.

A lo largo de su historia, Apple se ha acreditado como líder en el lanzamiento de productos que han revolucionado el mercado tecnológico, y que recientemente lanzó las Apple Vision Pro, y Meta se ha destacado por trabajar en el metaverso aunado al Oculus Insight – Realidad Virtual + Realidad Aumentada + IA -.

Las nuevas tecnologías que Apple ha introducido se definen como un producto innovador que se desmarca de las categorías convencionales en las que ya tiene presencia en el mercado: las Apple Vision Pro, que consiste en un dispositivo desarrollado

durante más de una década, anunciado en junio del año 2023.

Aunque puede ser percibido como una sofisticada diadema de comunicación, Apple ha optado por llamarlo “computadora espacial” debido a su capacidad para fusionar el mundo digital con el físico.

Este dispositivo representa una incursión significativa en la realidad mixta, pues ofrece contenido tanto de realidad aumentada como de realidad virtual inmersiva. Es necesario destacar que todo lo visualizado en el dispositivo es digital, ya que utiliza cámaras para mapear y procesar el entorno circundante. [1]

De otro lado, la tecnología Oculus Insight, privilegió la eficiencia del procesamiento que se logró mediante la optimización de los algoritmos de visión por computadora para funcionar en hardware móvil, proceso en el que utilizó técnicas como la paralelización de operaciones y la optimización del uso de memoria, lo cual implicó un diseño cuidadoso de los algoritmos para minimizar la cantidad de cálculos necesarios sin comprometer la precisión del seguimiento.

Además, estas tecnologías han utilizado técnicas de compresión de datos y transmisión de información entre los diferentes componentes del sistema para optimizar los procesos de transferir y procesar, lo que ha ayudado a reducir la carga computacional y el consumo de energía.

En resumen, estas tecnologías de vanguardia han requerido un enfoque multidisciplinario que ha combinado investigación avanzada en visión por computadora, algoritmos de seguimiento de movimiento, optimización de hardware y técnicas de simulación para garantizar un seguimiento preciso y eficiente en dispositivos de realidad virtual todo en uno.

Esta combinación de tecnologías y enfoques innovadores han permitido ofrecer experiencias de realidad virtual verdaderamente inmersivas, libres de cables y comandada mediante voz, ojos y gestos.

2. DISEÑO

En este apartado se dan especificaciones de algunos de los productos mencionados en la introducción de este escrito; sin embargo, debe entenderse que son el estándar de la industria y que diversas empresas distribuirán sus productos en los términos técnicos referenciados.

Figura 1.
Imagen real de Apple Vision Pro [1]



Figura 2.
Meta Quest Pro [2]

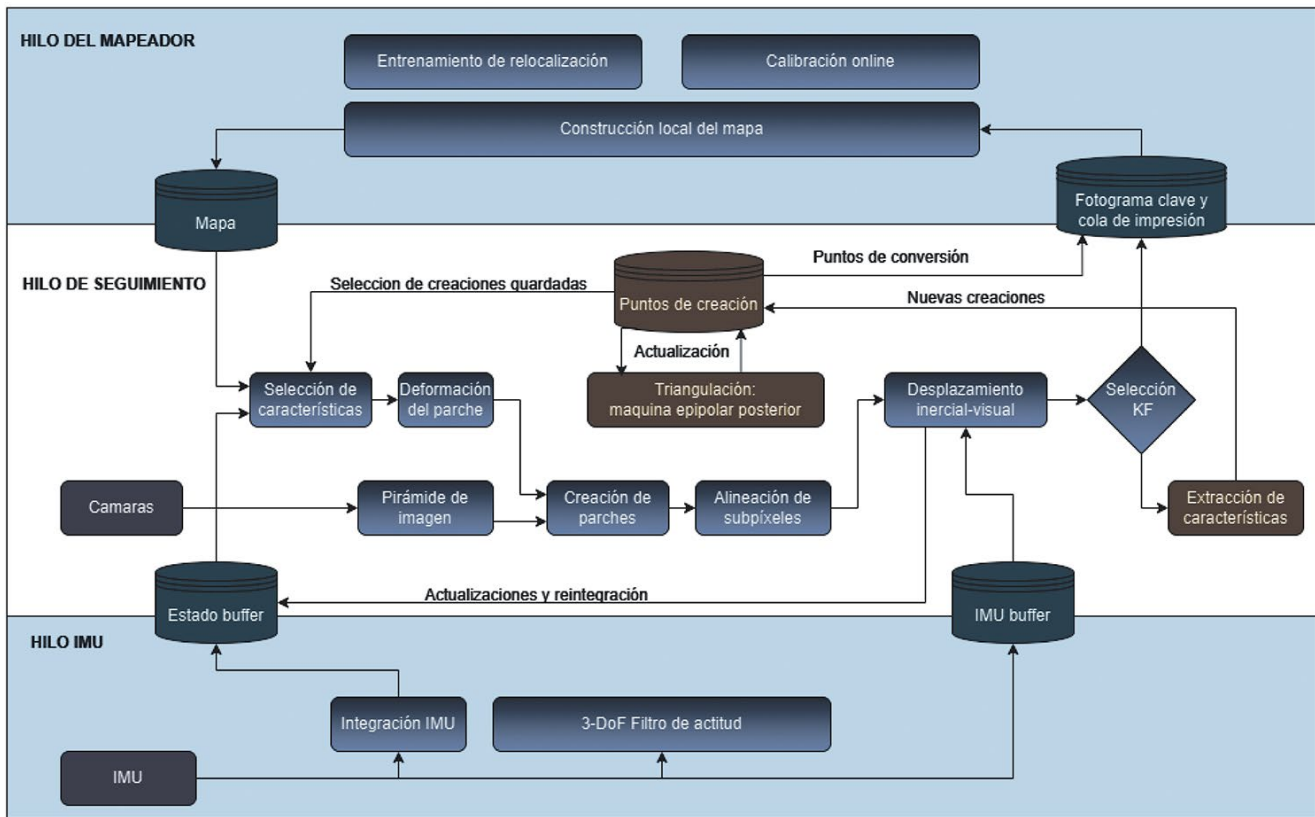


El diseño exterior es centrado en el usuario con un estilo vanguardista y futurista que brinda inmersión completa – realidad mixta - en todas actividades de la vida diaria, gracias al soporte del hardware de última generación que cuenta con especificaciones similares a estas:

Característica	Especificación	Descripción
Capacidad	256GB, 512GB, 1TB.	Necesaria para ejecutar los procesos y aplicaciones.
Display	23 millones de píxeles, sistema de visualización 3D, Micro-OLED	Admite reproducción múltiple de 24 fps. y 30 fps.
Chips	CPU de 8 núcleos con 4 núcleos de rendimiento y 4 núcleos de eficiencia, GPU de 10 núcleos	Memoria unificada de 16GB
RTDP	Procesamiento de datos en tiempo real	Procesadores dedicados a la gestión de imágenes en tiempo real. Latencia de fotón a fotón de 12 milisegundos, ancho de banda de memoria de 256 GB/s
Cámara	Sistema de cámara principal estereoscópica 3D	Captura espacial de fotografías y vídeos, 18 mm, apertura $f/2,00$, 6,5 megapíxeles estéreo.
Sensores	Cámaras de alta resolución, sensor de parpadeo, sensor de luz ambiental	Dos cámaras principales de alta resolución, seis cámaras de seguimiento orientadas al mundo, cuatro cámaras de seguimiento ocular, sensor de parpadeo, sensor de luz ambiental
Identificación óptica	Autenticación biométrica basada en iris	Generalmente se usa para temas de seguridad cibernética.
Tecnología de audio	Audio espacial con seguimiento dinámico de la cabeza. Matriz de seis micrófonos con formación de haz direccional	Los formatos admitidos incluyen AAC, MP3, Apple Lossless, FLAC, Dolby Digital, Dolby Digital Plus y Dolby Atmos. Los formatos admitidos incluyen HEVC, MV-HEVC, H.264, HDR con Dolby Vision, HDR10 y HLG
Batería	Hasta 2 horas de uso general, visualización de vídeo hasta 2,5 horas	Se puede utilizar mientras se carga la batería
Conectividad e inalámbrico	Wi-Fi 6 (802.11ax) y bluetooth 5.3	Se actualizan según los estándares del mercado para garantizar navegación y funcionamiento
Sistema operativo	SO de interfaz de usuario tridimensional controlado por ojos, voz y gestos	Depende del desarrollo de cada fabricante, se acoplará al dispositivo
Entradas	Manos, ojos y voz	Accesorios de entrada compatibles: Teclados, paneles táctiles y controladores de juego
Peso	600 a 650 g (21,2 a 22,9 onzas)	La batería separada pesa 353 g.

Este conjunto de características técnicas del dispositivo que utiliza el usuario se in-

serta en una arquitectura del sistema que se refleja en el siguiente diagrama:



3. FUNCIONAMIENTO

Este tipo de dispositivos proveen al usuario la autenticación mediante el reconocimiento de iris, de tal forma que solo funcionan cuando el usuario registrado las utiliza. Estas cámaras infrarrojas escanean el iris, a cambio del escaneo de cara o escaneo de huella, a lo cual se denomina capa de seguridad de autenticación en el nivel de aplicaciones y es válido en todo el ecosistema (VisionOS [1]). La seguridad usa un cifrado de primer nivel y todo lo que se realiza en el dispositivo solo se almacena allí y no se transmite a internet o a la nube ni a otros dispositivos sin autorización del usuario, es decir: los datos que se recopilan siempre están procesándose en el mismo dispositivo, y la información que se envía a las aplicacio-

nes o sitios web es solo la que elige el usuario final [4].

Las cámaras del iris también están encargadas de realizar un seguimiento ocular que rastrea el movimiento del ojo y detectar qué objeto se mira; es posible navegar en menús, abrir aplicaciones y controlar objetos con los ojos [1].

Otra gran función es el seguimiento manual que permite detectar cuando el usuario está realizando determinados movimientos de las manos, lo cual se hace a través de las cámaras; con esto se puede mover objetos, rotarlos y cambiar el tamaño (Figura 2). A la vez que va detectando el movimiento de las manos rastrea rotación de cabeza y desplazamiento para mostrar la actividad en pantalla [3].

Figura 3.

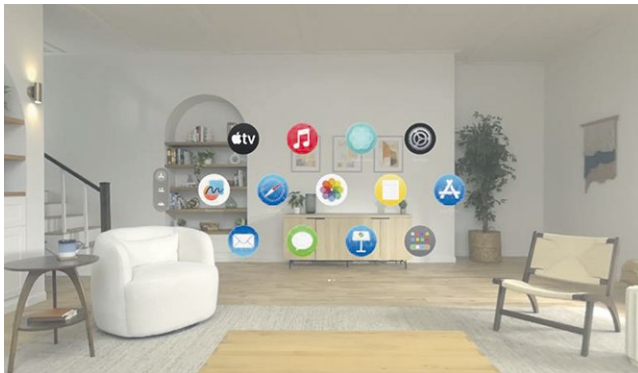
Gestos de la mano para navegación [3].



Por último, al considerarlo como un computador espacial, se entiende que cuenta con la capacidad de comprender la ubicación y el tamaño de los objetos del mundo real, lo que permite experimentar la realidad mixta que combina lo real con lo virtual y hacerlo uno solo para lograr la experiencia total (Figura 3). [3].

Figura 4.

*Así veríamos la experiencia [5].
Tomado de Apple Visión Pro.*



4. USOS

Actualmente, el alcance de toda nueva tecnología lanzada es amplio e imaginativo, ya que al introducirse como novedad puede generar experiencias inéditas que, en ocasiones, contribuyen al bienestar gene-

ral, reducción de costos, mejoras en la salud, fortalecimiento de la seguridad, entre otros aspectos.

En medicina se usa para cirugías de forma remota, y también para el entrenamiento de médicos y para brindar terapias [6]. Adicionalmente, se habla de poder mejorar la agudeza visual de la población, gracias a que la población que está localizada en zonas apartadas podrá acceder fácilmente a exámenes rutinarios o especializados que permitirán detectar de forma temprana enfermedades que causan pérdida de la visión [7].

En la industria, el uso de visores con interfaz de usuario tridimensional se ha extendido a diversas aplicaciones, como el diseño de productos, pruebas de prototipos, capacitación de trabajadores e incluso la visualización de datos para la toma de decisiones. Por ejemplo, el dispositivo Apple Vision Pro incluye una demostración que simula el funcionamiento de una turbina [3], lo cual permite a los usuarios interactuar y desmontar piezas (Figura 4).

Fig. 4

*Imagen de una turbina [8].
Tomado de Apple Visión Pro*



En educación se está usando para crear experiencia de aprendizaje inmersivo, recrear eventos históricos, actividades colaborativas interactivas, educación a distancia [6]; como ejemplo, la Universidad Libre tiene planteado crear la visualización de datos captados por satélites en un cuarto de realidad virtual [9], que permitirá su análisis e interacción.

5. CONCLUSIONES

En conclusión, la introducción de tecnologías innovadoras como las Apple Vision Pro y Oculus Insight en el campo de la realidad virtual y mixta representa un hito significativo en la evolución tecnológica. Estos dispositivos no solo han logrado fusionar de manera impresionante lo digital y lo físico, sino que también han demostrado un enfoque multidisciplinario al combinar avanzadas investigaciones en visión por computadora, algoritmos de seguimiento de movimiento, optimización de hard-

ware y técnicas de simulación. La autonomía, autenticación biométrica basada en iris, seguimiento ocular y manual, junto con la capacidad de comprender el entorno real, han catapultado estas tecnologías hacia experiencias de realidad virtual inmersivas y libres de cables, controladas mediante voz, ojos y gestos.

Estas gafas de RV no solo se destacan en el ámbito del entretenimiento, pues también tienen aplicaciones prácticas en medicina, industria y educación, lo cual abre nuevas posibilidades para cirugías remotas, entrenamiento médico, mejoras en la agudeza visual, diseño de productos, simulación de procesos industriales y experiencias de aprendizaje inmersivas. La realidad virtual está cada vez más arraigada en nuestro día a día, debido a que promueve innovaciones que trascienden los límites tradicionales y mejoran significativamente diversas áreas de la vida.

6. REFERENCIAS BIBLIOGRÁFICAS

- [1] M. Staff, «MacRumors,» 2024. [En línea]. Available: https://www.macrumors.com/roundup/apple-vision-pro/#apple_vision_pro.
- [2] «insight,» [En línea]. Available: <https://www.insight.com/>.
- [3] «Apple,» 2024. [En línea]. Available: <https://www.apple.com/apple-vision-pro/specs/>.
- [4] E. Feo, «Apple Vision Pro: The Future of Tech with a Wink and a Nod» 2024. [En línea]. Available: <https://bootcamp.uxdesign.cc/apple-vision-pro-the-future-of-tech-with-a-wink-and-a-nod-56c6a07a5153>.
- [5] Apple, «Support Apple» 2024. [En línea]. Available: <https://support.apple.com/guide/apple-vision-pro/apple-vision-pro-and-visionos-overview-tan39b6ba-b8f/1.0/visionos/1.0>.
- [6] ismail, «Apple Vision Pro: A Portal to a New Reality,» 2024. [En línea]. Available: <https://medium.com/@ismailsaleem/apple-vision-pro-a-portal-to-a-new-reality-b16a2050c2fd>.
- [7] E. Waisberg, J. Ong, M. Masalkhi, N. Zaman, P. Sarker, A. G. Lee y A. Tavakkoli, «The future of ophthalmology and vision science with the Apple Vision Pro,» Eye, vol. 38, pp. 242-243, 2023.
- [8] A. Talk, «Bootcamp,» 2024. [En línea]. Available: <https://bootcamp.uxdesign.cc/discover-the-best-apple-vision-pro-apps-a-starter-kit-6e946826545e>.
- [9] J. F. Robayo Ramirez, F. A. Simanca Herrera y R. A. Santa Quintero, «Visualización de datos captados por satélites en un showroom de realidad virtual,» Revista Avenir, vol. 7, pp. 1-9, 2023.