

español



Modernisation of VET through
Collaboration with the Industry

Ivan Minárik
Gregor Rozinaj
Renata Rybárová
Marek Vančo
Radoslav Vargic

Nuevas Formas de Control de Sistemas



Erasmus+

El presente proyecto ha sido financiado con el apoyo de la Comisión Europea.
Esta publicación (comunicación) es responsabilidad exclusiva de su autor. La
Comisión no es responsable del uso que pueda hacerse de la información aquí
difundida.

Título: Nuevas Formas de Control de Sistemas

Autor: Ivan Minárik,
Gregor Rozinaj,
Renata Rybárová,
Marek Vančo,
Radoslav Vargic

Traducido por: Juan Antonio Ortega

Publicado por: České vysoké učení technické v Praze
Fakulta elektrotechnická

Dirección de contacto: Technická 2, Praha 6, Czech Republic

Número de teléfono: +420 224352084

Print: (only electronic form)

Número de páginas: 37

Edición: Primera edición, 2019

MoVET

Modernisation of VET through
Collaboration with the Industry

<https://movet.fel.cvut.cz>



El presente proyecto ha sido financiado con el apoyo de la Comisión Europea.

Esta publicación (comunicación) es responsabilidad exclusiva de su autor. La Comisión no es responsable del uso que pueda hacerse de la información aquí difundida.

NOTAS EXPLICATIVAS



Definición



Interesante



Nota



Ejemplo



Resumen



Ventajas



Desventajas

ANOTACIÓN

Las nuevas formas de control de sistemas están basadas en las últimas tecnologías y usan el hardware correspondiente. Los sistemas actuales pueden ser controlados sin hardware adicional como ratones o teclados. Los usuarios pueden controlar el sistema o navegar por la aplicación simplemente usando sus manos (navegación gestual), su voz (navegación por voz) o sus ojos (seguimiento ocular). Algunos sistemas pueden usar un interfaz cerebro-ordenador. Este módulo introduce las tecnologías citadas con el fin de ayudar a comprender sus principios básicos que nos encontramos en nuestra vida diaria.

OBJETIVOS

El objetivo principal del módulo es introducir al estudiante los conceptos fundamentales de las nuevas formas de control de sistemas en diferentes sistemas o aplicaciones. El estudiante estará claramente familiarizado con los principios básicos de navegación por gestos, navegación por comandos de voz, seguimiento ocular, interfaz cerebro-computadora y motor de recomendación.

LITERATURA

- [1] Vančo, Marek; Minárik, Ivan; Rybárová, Renata. Evolution of static gesture recognition. In: Redžúr 2014 proceedings; 8th International Workshop on Multimedia and Signal Processing; 13 May 2014, Dubrovnik, Croatia. Bratislava: Nakladateľstvo STU, 2014, p. 41-44. ISBN 978-80-227-4162-0.
- [2] Rozinaj, Gregor, et al. Extending System Capabilities with Multimodal Control. Acta Polytechnica Hungarica, 2016, 13.4.
- [3] <https://medium.com/iotforall/how-gesture-control-will-transform-our-devices-32d4527a6d25>
- [4] <http://www.thedrive.com/aerial/10674/djis-spark-is-a-hand-gesture-controlled-drone-that-flies-off-your-hand>
- [5] <https://stfalcon.com/en/blog/post/intuitive-gestures-in-mobile-app-design>
- [6] Parrado Rollan, Marina; Posoldová, Alexandra; Rybárová, Renata. Recommendation engine design using Bayesian network for feature inference. In Redžúr 2016, 10th International workshop on multimedia and signal processing. Bratislava, Slovakia. May 24, 2016. 1. ed. Bratislava: Nakladateľstvo STU, 2016, CD-ROM, pp. 57-60. ISBN 978-80-227-4560-4

- [7] Grau, C., Ginhoux, R., Riera, A., Nguyen, T. L., Chauvat, H., Berg, M., ... Ruffini, G. (2014). Conscious Brain-to-Brain Communication in Humans Using Non-Invasive Technologies. PLoS ONE, 9(8), e105225. <http://doi.org/10.1371/journal.pone.0105225>
- [8] Guy V. et al. Brain computer interface with the P300 speller: Usability for disabled people with amyotrophic lateral sclerosis, 2018, Annals of Physical and Rehabilitation Medicine, 61 (1) , pp. 5-11.
- [9] Fernández, R. et al. Review of real brain-controlled wheelchairs, 2016, J. Neural Eng. 13 061001, <https://doi.org/10.1088/1741-2560/13/6/061001>
- [10] Kosmyna, N., Tarpin-Bernard, F., Bonnefond, N., & Rivet, B. (2016). Feasibility of BCI Control in a Realistic Smart Home Environment. Frontiers in Human Neuroscience, 10, 416. <http://doi.org/10.3389/fnhum.2016.00416>
- [11] Furness, D., The University of Florida just held the world's first mind-controlled drone race, available online: <https://www.digitaltrends.com/cool-tech/mind-controlled-drone-race-university-of-florida/>
- [12] Chen,S., “Forget the Facebook leak’: China is mining data directly from workers’ brains on an industrial scale”, available online: <https://www.scmp.com/news/china/society/article/2143899/forget-facebook-leak-china-mining-data-directly-workers-brains>
- [13] Chen, A., “Brain-scanning in Chinese factories probably doesn’t work — if it’s happening at all”, published 1.5.2018, available online: <https://www.theverge.com/2018/5/1/17306604/china-brain-surveillance-workers-hats-data-eeg-neuroscience>
- [14] Martišius, I., Damaševičius, R., “A Prototype SSVEP Based Real Time BCI Gaming System,” Computational Intelligence and Neuroscience, vol. 2016, Article ID 3861425, 15 pages, 2016. <https://doi.org/10.1155/2016/3861425>.
- [15] Narayanan, A et al., “Toward domain-invariant speech recognition via large scale training,” Google, USA, arXiv:1808.05312, 2018. <https://arxiv.org/abs/1808.05312>

Indice

1	Introducción al control de sistemas	7
2	Tecnologías para el control del sistema.....	8
3	Diferentes tipos de control de sistemas.....	16
3.1	Control del sistema mediante gestos	16
3.2	Control del sistema mediante comandos de voz	18
3.3	Control del sistema mediante seguimiento ocular.....	20
3.4	Control del sistema vía interfaz cerebro-computadora (BCI)	21
3.5	Control del sistema vía motor de recomendaciones	23
4	Control del sistema en aplicaciones	25
4.1	Comandos de voz para dispositivos móviles.....	25
4.2	Gestos para sistemas modernos de control de TV	27
4.3	Gestos para teléfonos inteligentes (y otras aplicaciones).....	31
4.4	Seguimiento ocular (control del ratón del PC, navegación entre opciones en pantalla)	34
4.5	BCI (control de sillas de ruedas, navegación en juegos).....	35
5	¿Conclusión o quo vadis, control del sistema?	37

1 1 Introducción al control de sistemas

Hoy en día, las tecnologías de la información se están poniendo cada vez más en primer plano en nuestra vida cotidiana. La forma de administrar o controlar dispositivos y sistemas es cada vez más cómoda y fácil de usar. En las siguientes páginas se introducirán nuevas y modernas tecnologías para el control del sistema y la navegación.

Hoy en día, los gestos son una forma muy popular para el control de aplicaciones o sistemas y muchas personas los utilizan todos los días. Los controles gestuales son naturales porque están asociados con la forma en que las personas interactúan con objetos reales. En realidad, estamos utilizando gestos en nuestros dispositivos móviles, aplicaciones informáticas, aplicaciones de *realidad aumentada/realidad virtual*, (**AR/VR**, siglas del inglés augmented reality/virtual reality), consolas de videojuegos, etc. Gracias a su uso diario, las interfaces gestuales comienzan a implantarse en otras áreas de la tecnología. Se espera que la interacción gestual esté disponible en casi todos los dispositivos en unos pocos años. La alta popularidad de la navegación gestual obliga a los investigadores a mejorar estas tecnologías. Esta es una tendencia obvia, ya que el rendimiento de los equipos informáticos ya no es el cuello de botella para una navegación y un control más naturales mediante gestos [1].

Debido a la existencia de la gran cantidad de información disponible a solo un clic de distancia, no es posible que una persona se mantenga al día con todos los datos en los que pueda estar interesada. Mediante el uso de la tecnología, la información se puede seleccionar de antemano para los usuarios. Por eso, un Motor de Recomendación es una potente herramienta que puede hacer que los usuarios dejen de perder el tiempo buscando y revisando información que realmente no les importa. Basándose en sus preferencias, el sistema puede predecir lo que será adecuado para el usuario [6].

En los últimos años, el reconocimiento de voz ha dado un salto considerable tanto en velocidad como en calidad.

Una tecnología muy interesante para el control de dispositivos es el *Interfaz Cerebro-Computadora* (**BCI**, siglas del inglés Brain Computer Interface). BCI representa (expresándolo de forma simple) una vía de comunicación directa entre el cerebro humano y el dispositivo externo. En este módulo introduciremos también el seguimiento ocular, donde se utiliza la medición de la actividad ocular para controlar el dispositivo.

2 **2** Tecnologías para el control del sistema

Las tecnologías utilizadas para el control del sistema pueden diferir según el área donde se usan. Para diferentes tipos de control del sistema (navegación gestual, navegación por voz, seguimiento ocular, etc.), también se utilizan hardwares diferentes.

HW usado en *el proceso de reconocimiento de gestos*:

- Dispositivos de Control Gestual – en el pasado se usaron guantes cableados para capturar gestos y movimientos de las manos. Los guantes utilizan interruptores táctiles, sensores ópticos o resistivos para medir la flexión de las articulaciones.
- Reconocimiento Gestual Basada en Visión - utiliza una cámara genérica y / o una cámara de alcance (range camera en inglés) para capturar e identificar el gesto de la mano. Existen múltiples métodos para el reconocimiento de gestos basados en el uso de cámaras.
- Cámaras 3D - pueden percibir profundidad. Las cámaras 3D se han vuelto mucho más accesibles y baratas en los últimos años [3].

Pantallas táctiles

Generalmente, podemos distinguir dos tipos de pantallas táctiles: *resistivas* y *capacitivas*.

$E=mc^2$

Una pantalla táctil resistiva consta de varias capas, de las cuales las capas flexibles de plástico y vidrio son dos capas importantes de resistencia eléctrica.

Ambas capas se enfrentan entre sí y entre ellas hay un espacio fino. Cuando un dedo o la punta de un lápiz presiona la superficie exterior, ambas películas se encuentran. La medida de la resistencia de ambas capas en el lugar de contacto permite una medición precisa de la posición táctil.

+

Ventajas de una Pantalla Táctil Resistiva:

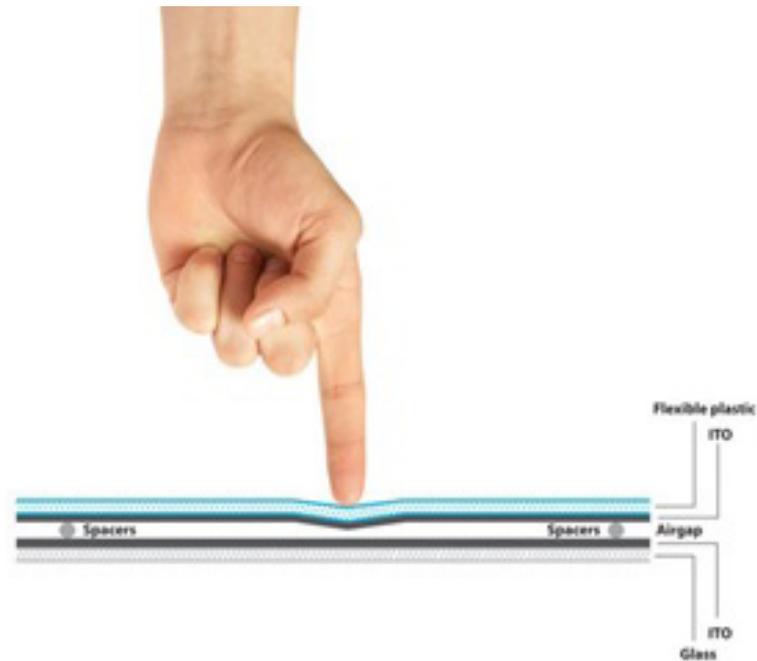
- Alta resistencia al polvo y al agua
 - Se utiliza mejor con un dedo, mano enguantada o lápiz óptico
 - Mejor adecuada para el reconocimiento de escritura a mano
-

-

Desventajas de una Pantalla Táctil Resistiva:

- No demasiado sensible, se debe apretar fuerte
- Pobre contraste debido a tener reflexiones adicionales de la capa extra de material colocada sobre la pantalla

- No soporta multi-toque
-



Pantalla táctil resistiva

Pantalla Táctil Resistiva

$E=m \cdot c^2$

Una pantalla táctil capacitiva consiste también en dos capas separadas de vidrio, que están recubiertas con un conductor como Óxido de Estaño de Indio (ITO).

El cuerpo humano es un conductor de carga eléctrica. Cuando un dedo toca el vidrio de la superficie capacitiva, cambia el campo electrostático local. El sistema monitoriza continuamente el movimiento sobre cada pequeño condensador para averiguar el área exacta donde el dedo ha tocado la pantalla.

+

Ventajas de una Pantalla Táctil Capacitiva:

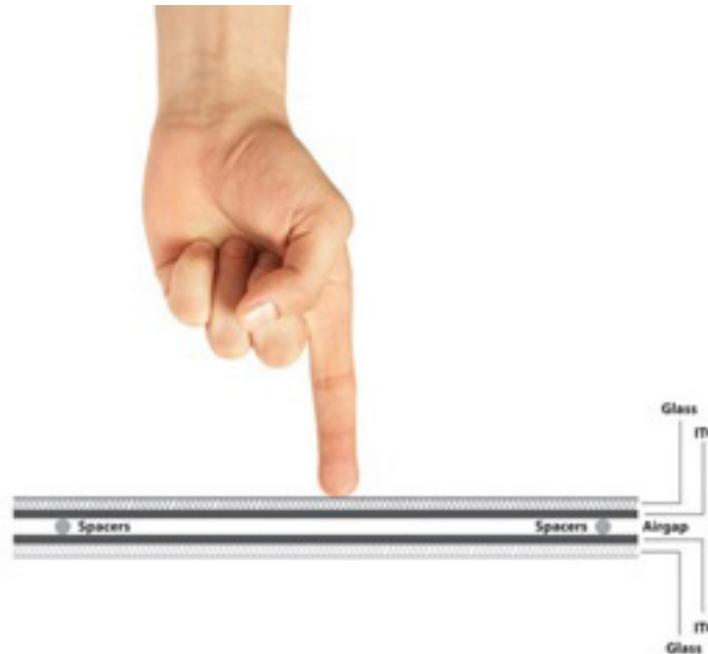
- Debido a que la pantalla táctil capacitiva tiene una capa de vidrio en lugar de plástico, se ve más brillante y nítida
 - Muy sensible al tacto y no necesita un lápiz óptico
 - Soporta multi-toque
-

-

Desventajas de una Pantalla Táctil Capacitiva:

- Debido a que la tecnología depende de la naturaleza conductora del cuerpo humano, no funciona si el usuario usa guantes

- Debido a que tienen una estructura compleja, son bastante caras
 - El vidrio es más propenso a romperse
-



Pantalla táctil capacitiva

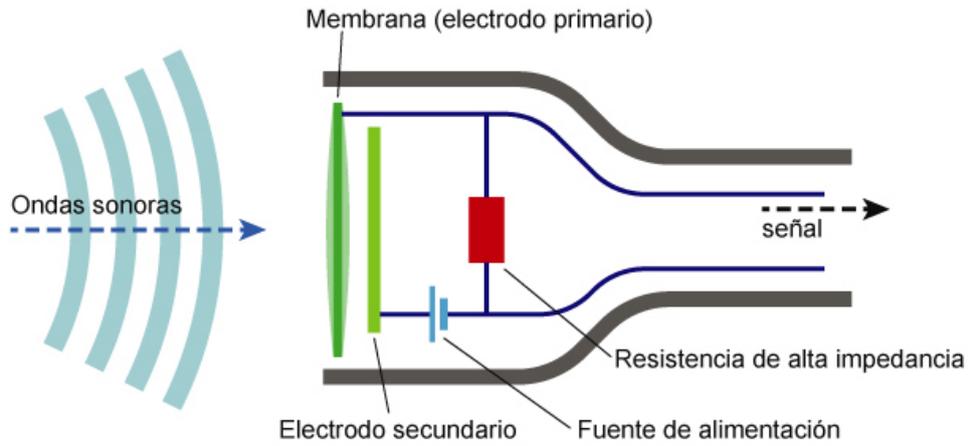
Pantalla táctil capacitiva

Micrófonos

Un micrófono convierte ondas acústicas en señales eléctricas. Un diafragma reacciona vibrando a las ondas acústicas, y produce cargas eléctricas de la intensidad correspondiente. Existen varias técnicas para generar la señal eléctrica, por ejemplo, condensador, dinámico, piezoeléctrico o incluso láser. Los teléfonos móviles suelen utilizar micrófonos electret, o **MEMS** (*Sistemas Micro Eléctrico Mecánicos o MicroElectrical-Mechanical System en inglés*).

Micrófono de Condensador

Dos placas se alimentan para crear un condensador. Una de las placas actúa como una membrana y se mueve en base a las ondas acústicas entrantes. El movimiento cambia la tensión de salida, generando la señal.



Esquema principal de un micrófono de condensador



Por lo general, un micrófono de estudio, es más sensible que el micrófono dinámico. Se utiliza para grabar instrumentos musicales.



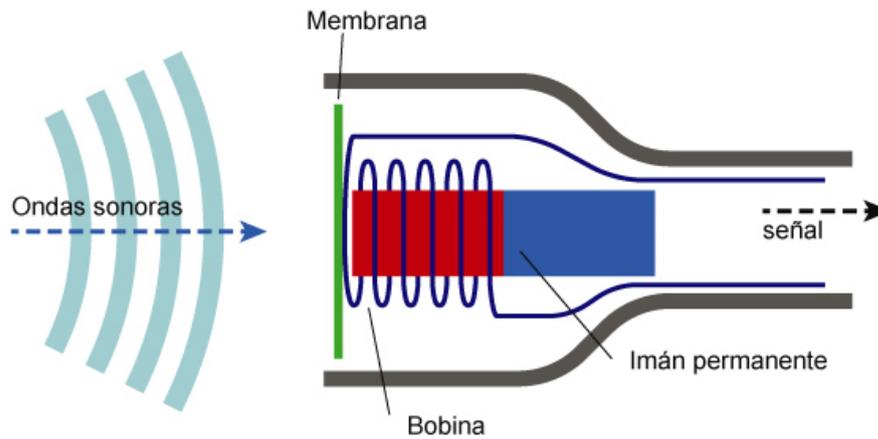
El micrófono necesita una fuente de alimentación.



Un micrófono electret es una actualización tecnológica del micrófono de condensador, haciéndolo más resistente. El micrófono electret se usa en la mayoría de los dispositivos móviles en la actualidad.

Micrófono Dinámico

La membrana del micrófono se conecta a una bobina ubicada alrededor de un imán permanente. La presión aplicada a la membrana obliga a la bobina a moverse a lo largo del imán y genera corriente eléctrica.



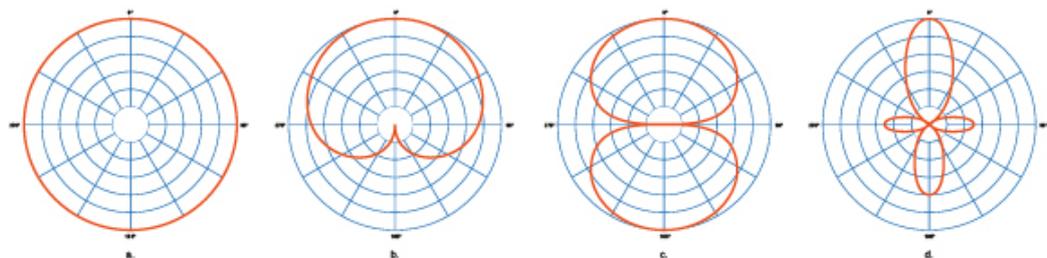
Esquema principal de un micrófono dinámico



El micrófono dinámico es menos sensible, lo que lo hace más adecuado para grabar canciones en vivo.

Otras tecnologías para la adquisición de sonido incluyen carbón, piezoeléctrico, cinta, MEMS, líquido o láser.

Según la forma de los componentes básicos, un micrófono puede tener una sensibilidad diferente a los diferentes ángulos de la fuente de sonido. El más común es el patrón cardioide, que hace que el micrófono capte las ondas de sonido que se encuentran frente a él, pero que no lo haga por detrás. Otros patrones de sensibilidad incluyen omnidireccional, bidireccional y direccional.



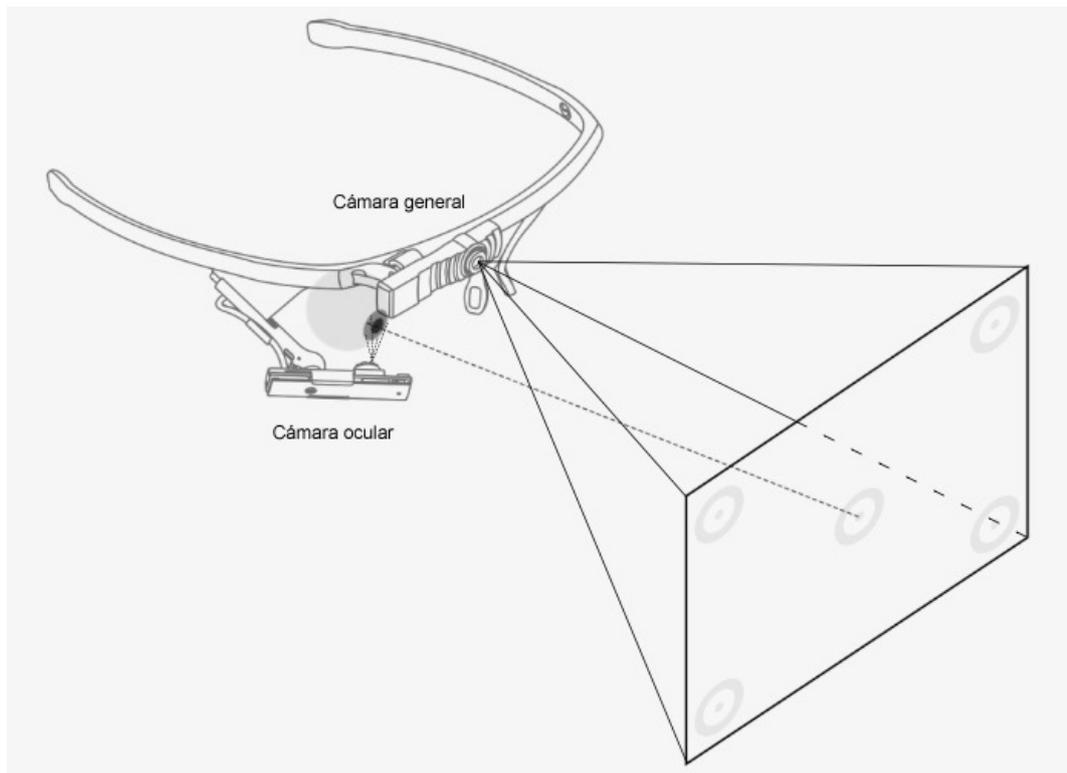
Muestra de las Características Direccionales de un Micrófono. A. Omnidireccional. B. Cardioide. C. Bidireccional (figura de 8). D. Direccional (Escopeta).

Una combinación de varios micrófonos (por ejemplo, ordenados en una fila) crea una matriz de micrófonos. La matriz puede enfocarse mejor hacia el área seleccionada mientras ignora el área innecesaria. Una característica importante de las matrices de micrófonos es la capacidad de deducir la dirección de la que proviene la voz del orador. Esto ayuda a apuntar a otros sistemas del sistema multimedia.

Seguimiento ocular y BCI

El seguimiento ocular es el proceso de medir la posición del punto de la mirada (donde está mirando la persona rastreada) o el movimiento de los ojos en relación con la cabeza.

Los sistemas para el seguimiento ocular se basan principalmente en cámaras que capturan el ojo o las imágenes de los ojos y en base a esto evalúan la posición de la mirada. Las cámaras capturan la imagen del ojo normalmente con una frecuencia de cuadro desde 30Hz (nivel de entrada, juego) hasta 1200Hz (nivel de investigación). Básicamente, existen dos construcciones de rastreadores oculares: la variante móvil, donde las cámaras que capturan la posición de los ojos están montadas en gafas o incorporadas en un HMD, o la variante fija, donde las cámaras se colocan en una caja debajo de la pantalla / monitor. Las versiones binoculares móviles (cada ojo se rastrea mediante una cámara dedicada) de seguidor ocular son generalmente más precisas que las monoculares (solo se rastrea un ojo) y permiten movimientos oculares más grandes. Además de las cámaras oculares, en la variante móvil también se utiliza la cámara “general”, que captura el entorno y permite trazar la posición de la mirada en la imagen circundante.

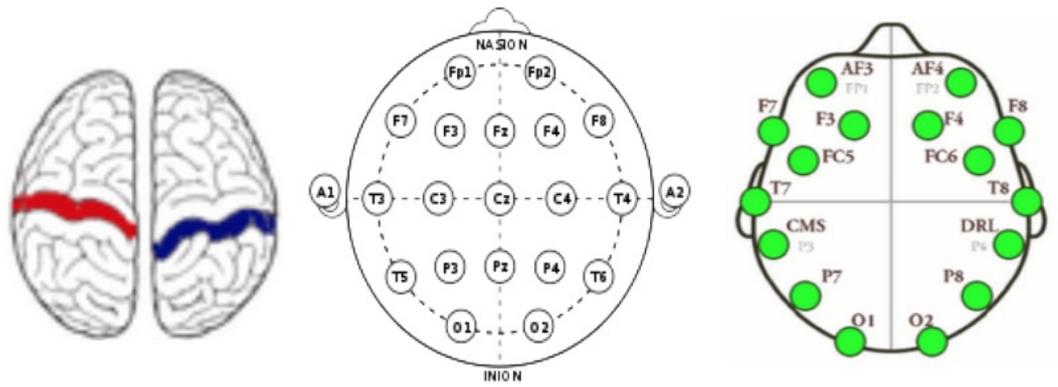


Principio del rastreador de ojo monocular móvil, el visor es escaneado por la cámara general y la posición de la pupila por la cámara ocular

Básicamente, existen dos construcciones de rastreadores oculares: la variante móvil, donde las cámaras están montadas en gafas o incorporadas en un HMD, o la variante fija, donde las cámaras se colocan en una caja debajo de la pantalla / monitor. Además de las cámaras “oculares” mencionadas anteriormente, en la variante móvil también se utiliza la cámara “general”, que captura el entorno y permite trazar la posición de la mirada en la imagen circundante.

Una Interfaz Cerebro-Computadora (BCI) es una tecnología que permite la comunicación entre un cerebro humano y un sistema externo (generalmente basado en computadora). BCI puede referirse a una tecnología que lee señales del cerebro por un sistema externo y/o a una tecnología que envía señales al cerebro.

Para el control del dispositivo, el interés principal es leer las señales del cerebro e interpretar la intención del usuario. Las señales enviadas al cerebro se pueden utilizar como canal de retroalimentación. La tecnología BCI para enviar las señales al cerebro puede usar, por ejemplo, *estimulación magnética* (TMS del inglés transcranial magnetic stimulation) [7]. El TMS es un enfoque no invasivo en el que se utiliza un campo magnético cambiante para generar una corriente eléctrica en la región objetivo del cerebro a través de inducción electromagnética. La tecnología BCI para leer las señales cerebrales en su mayoría utiliza *electroencefalografía* (EEG) – señales eléctricas creadas por las neuronas y capturadas en la piel sobre el cráneo mediante electrodos, que suelen estar chapados en oro o húmedos. Normalmente, los sistemas BCI usan 2 electrodos (nivel de entrada, juego) - 4 y más, hasta 256 (nivel de investigación). La parte importante de las señales capturadas (ondas cerebrales) se encuentran en la banda de frecuencia de 2Hz-30Hz, son muy débiles (2-30mV) y necesitan ser amplificadas. Este rango de frecuencia se divide entre más subbandas (también llamadas ondas cerebrales), como beta, theta, etc. La presencia de energía en estas subbandas puede indicar varias situaciones. Depende también de la ubicación de la medición. Por ejemplo, las ondas delta hasta 4 Hz, ubicadas en la parte frontal y para los adultos, están presentes en más etapas del sueño. Las ondas alfa entre 7 Hz y 13 Hz ubicadas en las regiones posteriores de la cabeza, a ambos lados, aumentan con el cierre del ojo y con la relajación, y se atenúan con la apertura del ojo o el esfuerzo mental. Las ondas mu entre 8Hz y 13Hz, están ubicadas en la corteza sensoriomotora (parte central superior en el cuero cabelludo en ambos lados) y están presentes durante las acciones motoras o incluso en la imaginación de las acciones motoras. Un problema es que también se capturan las señales mio (señales generadas debido a los movimientos musculares) y éstas son un orden de magnitud más grandes (10-300 mV). Por lo tanto, es necesario un postproceso cuidadoso. También hay sistemas que se basan en señales mio, capturadas por ejemplo en la muñeca o en el antebrazo, para captar los gestos. En su mayoría, los sistemas basados en señales mio realmente están pensados para ser usados en soluciones neuroprotésicas. Un caso especial es la oculografía (que captura las señales de los músculos del movimiento del ojo). Estos sistemas ahora son ampliamente reemplazados por los sistemas de seguimiento ocular basados en cámaras mencionados anteriormente.



Ubicaciones espaciales relevantes BCI para la adquisición de la señal mediante EEG: a) Ubicación del córtex SMR (rojo - parte motora, azul - sensorial) b) Sistema 10/20 para colocación del electrodo c) Colocación del electrodo del dispositivo Emotiv EPOC BCI

3 **3** Diferentes tipos de control de sistemas

3.1 Control del sistema mediante gestos

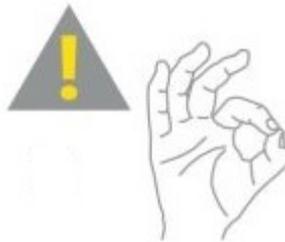
Actualmente, los dispositivos de entrada más utilizados para la comunicación entre personas y computadoras son el teclado, el ratón o la tableta táctil. Estos dispositivos están lejos de ser el ideal de una comunicación natural con una computadora, y más bien representan una adaptación humana a las limitaciones de la computadora. En los últimos años, comenzó a surgir el requisito de que los humanos necesitan comunicarse con las máquinas de la misma manera que lo hacen entre sí: mediante el habla o los gestos, ya que generan mucha más información que el enfoque de los dispositivos periféricos.

Los gestos se interpretan de forma natural en nuestros teléfonos inteligentes, tabletas, computadoras, etc. Su misión es hacer que la comunicación entre humanos y computadoras y su control sean fácil. Los gestos pueden ser táctiles o no táctiles, pero los principios fundamentales siguen siendo los mismos.

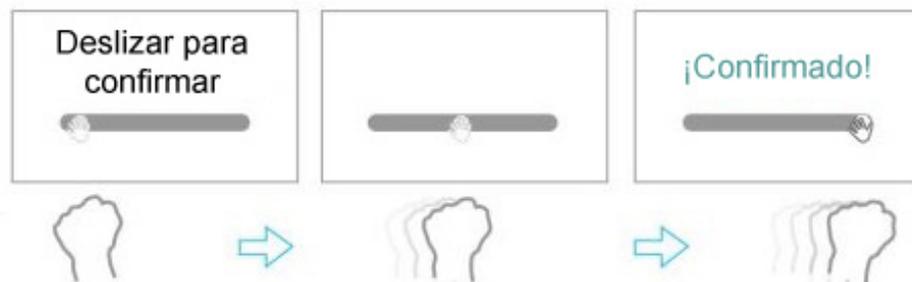


Los gestos se pueden dividir en dos categorías básicas según la experiencia del usuario.

- Los gestos innatos se basan en la experiencia general de todos los usuarios, como mover un objeto hacia la derecha moviendo la mano hacia la derecha, atrapar un objeto con los dedos cerrados, etc. Naturalmente, los gestos innatos pueden verse afectados por los hábitos o la cultura. Con los gestos innatos no hay necesidad de que un usuario los estudie para obtener una buena experiencia de control, solo necesitan que se los muestren.
 - La segunda categoría son los gestos aprendidos, los cuales es necesario aprender. Los gestos también se pueden dividir en tres categorías según la noción de movimiento.
 - Los gestos estáticos representan formas creadas por los miembros gesticulantes, que generan una información significativa. El reconocimiento de cada gesto es ambiguo debido a la oclusión de la forma de la extremidad y, en el nivel más alto de reconocimiento, el significado real del gesto basado en las propiedades culturales locales [1].
 - En la segunda categoría, los gestos continuos sirven como base para una interacción con la aplicación donde no se reconoce una postura específica, pero un movimiento solo se usa para reconocer el significado de un gesto [1].
 - Los gestos dinámicos consisten en un movimiento específico y predefinido de la extremidad con la que se realiza. Dicho gesto se usa para manipular un objeto o para enviar un comando de control [1].
-



Gesto estático



Gesto Dinámico



El uso de gestos para la navegación y el control del sistema proporcionará la Interfaz de usuario natural (NUI, del inglés Natural User Interface), que elimina completamente la dependencia de cualquier dispositivo mecánico, como un teclado o un ratón. El contribuyente clave de NUI es el control de gestos sin contacto que permite manipular objetos virtuales de manera similar a los físicos. Los NUI permiten a los usuarios sumergirse rápidamente en el "nuevo mundo": aplicaciones con control maestro con mínimo aprendizaje, lo que es muy importante para las aplicaciones de AR / VR y los sistemas de inteligencia ambiental. En aplicaciones en expansión como el control autónomo de drones y la navegación de información y entretenimiento en el automóvil, NUI puede aumentar considerablemente la facilidad de uso [3].

3.2 Control del sistema mediante comandos de VOZ

El reconocimiento de voz representa una tendencia ascendente en la interacción con dispositivos de consumo [2]. La voz es la forma más natural de comunicación de persona a persona y contiene la mayoría de la información comunicada.



Los comandos de voz son una herramienta valiosa para controlar dispositivos y sistemas cuando los gestos o las interfaces táctiles no son adecuados. Su uso va desde los sistemas de entretenimiento en el hogar hasta el control de sistemas de información y entretenimiento (infotainment) de automóviles y el control para discapacitados físicos.



El reconocimiento de voz cubre varios sub-campos, a saber, la identificación del hablante y el reconocimiento de comandos de voz. En este último se centran los investigadores hoy en día gracias a los avances significativos en la tecnología de redes neuronales.

Generalmente, un sistema de reconocimiento de voz funciona en estos dos modos:

- Aprendizaje
- Reconocimiento

Durante el aprendizaje, el sistema aprende sobre todas las entradas posibles y su significado. Esto suele ocurrir en un dominio paramétrico; ya sean parámetros para comandos de voz individuales o información específica del hablante. Durante el reconocimiento, un patrón de entrada desconocido se empareja con la referencia más cercana de los patrones paramétricos aprendidos. Ambos pasos funcionan mejor con mayor calidad y cantidad de datos de entrada.



El reconocimiento de voz es propenso al reconocimiento incorrecto debido a la presencia de ruido u otros locutores que hablan simultáneamente.

Sin embargo, cuantos más datos tenga que procesar un sistema, más tiempo llevará. Y el tiempo es crucial cuando queremos lograr un reconocimiento de voz agradable y continuo.

Si miramos hacia atrás unos años, la mayoría de los sistemas de reconocimiento de voz permitieron reconocer solo un conjunto limitado de comandos aislados o a un locutor de una base de datos limitada. Esto llevaría a un conjunto de comandos altamente especializado.

Con la disponibilidad amplia y asequible de los servicios basados en la nube, los sistemas de reconocimiento de voz pueden hacer uso de soluciones rápidas basadas en servidor. Esto, combinado con la conexión a Internet de alta velocidad ampliamente disponible, permite que las interfaces de usuario actuales procesen

entradas de voz más complejas (generalmente, esto se aplica a cualquier patrón de señal de entrada). La combinación permite utilizar decisiones complejas tomadas por redes neuronales en el lado del servidor, lo que elimina la necesidad de una potente configuración de hardware y software del usuario. Además, las redes neuronales hacen que el reconocimiento de comandos aislados sea tan eficiente que ahora se pueden usar para reconocer comandos complejos compuestos por varios comandos o tipos de comandos.

El progreso en la utilización de redes neuronales sobre un hardware cada vez más potente permite mejoras en varias áreas. En primer lugar, el sistema se vuelve más independiente del entorno. Los parámetros de voz grave se distinguen en condiciones de audio cambiantes [15]. Además, el sistema es capaz de reconocer no solo palabras o frases específicas, sino también reconocer oraciones de frases completas, con matices y variaciones en las palabras utilizadas. Además, al incorporar la voz reconocida previamente, los sistemas pueden deducir el significado de la oración o comando, incluso si son vagos o inespecíficos. Los sistemas actualmente comienzan a comprender el contexto real en el que se reconoció el discurso y permiten reaccionar de manera más apropiada. Esto significa que los sistemas comienzan a comprender no solo el discurso real, sino también la idea oculta detrás de las palabras.

3.3 Control del sistema mediante seguimiento ocular

Los sistemas de seguimiento ocular evalúan la dirección de la mirada según la posición de la pupila. Utilizan las señales de vídeo de las cámaras para rastrear la posición de la pupila. Las señales de vídeo son procesadas por la unidad central, donde se evalúa la posición de la pupila y se estima la dirección de la mirada del usuario. Para transformar la posición de la pupila reconocida en la dirección de la mirada, se utiliza un modelo de ojo para un usuario particular. Los parámetros del modelo de ojo se estiman durante el proceso de calibración.

3.4 Control del sistema vía interfaz cerebro-computadora (BCI)

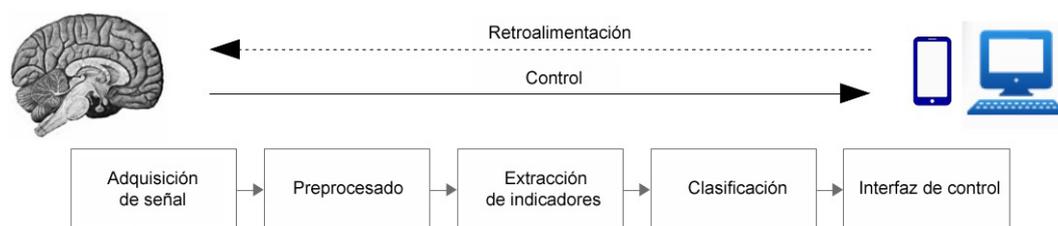
El control de sistemas utilizando BCI es uno de los desafíos para las opciones futuras de control de sistemas.



BCI tiene más ventajas como privacidad (sin sonidos fuertes, sin gestos visibles) y pocos cálculos (solo se capturan y procesan muy pocos datos comparándolo, por ejemplo, con un vídeo).



Pero BCI tiene en este momento también importantes inconvenientes como la falta de comodidad (se necesita una diadema, un auricular o un casco) y la necesidad de esfuerzo mental (cierto esfuerzo mental, en la mayoría de los casos, es necesario para "generar" las señales de control).



Un esquema conceptual de BCI

El esquema general del sistema BCI simple se muestra en la Fig. 1. La primera etapa es una adquisición de señal. El EEG mide la actividad cerebral eléctrica durante las excitaciones sinápticas en las neuronas. Las señales de EEG se capturan de forma no invasiva utilizando electrodos en el cuero cabelludo. Después de la adquisición de la señal, las señales deben ser preprocesadas. En general, las señales cerebrales adquiridas están contaminadas por el ruido y otros artefactos causados por señales biológicas o señales externas como la red eléctrica, etc. Después de obtener las señales "sin ruido" en la fase de mejora de la señal, las características esenciales de las señales cerebrales son extraídas. Los métodos de extracción de características más comunes utilizados con señales de EEG incluyen las transformadas ortogonales discretas. Una vez que se extraen las características adecuadas, debe haber un método que clasifique la señal en las clases deseadas. Existen muchas categorías de técnicas de clasificación como: generativas. (*Modelo Mixto de Gaussianas* – **GMM** del inglés Gaussian Mixture Model), discriminativos (*Redes Neuronales* – **NN** del inglés Neural networks, *Máquinas de Soporte Vectorial* – **SVM**, del inglés Support Vector machines), no paramétricos es decir basados en muestras (*K vecinos próximos* – **KNN** del inglés K nearest neighbour), etc. Cada método tiene sus pros y sus contras y, por lo tanto, debe seleccionarse en función de los requisitos de la aplicación. La etapa de interfaz de control de BCI utiliza la salida de clasificación como una señal de control. Los enfoques se pueden dividir en: endógeno (basado en la autorregulación de los ritmos y potenciales del cerebro sin estímulos externos) y exógeno (utiliza la actividad neuronal provocada en el

cerebro por un estímulo externo). Los métodos más utilizados incluyen *potenciales corticales lentos* (**SCP** del inglés slow cortical potentials), ritmos sensoriomotores, *potenciales evocados visuales* (**VEP** del inglés visual evoked potentials) *incluyendo VEPs de estado estacionario* (**SSVEPs**, del inglés Steady-State), y P300. Se pueden operar dispositivos o procesos particulares utilizando la interfaz de control. Los BCI que usan ritmos sensoriomotores usan las ondas cerebrales Mu presentes durante la imaginación de las acciones motoras (por ejemplo, el movimiento de la mano). Esto es BCI endógeno. Por otro lado, el SSVEP es BCI exógeno. SSVEP se basa en la propiedad de que el córtex visual "resonó" de acuerdo con la frecuencia del estímulo visual que observa el sujeto. Por lo tanto, el uso de este método es necesario, por ejemplo, en los botones de control de movimiento de la pantalla, cada uno parpadea con una frecuencia diferente. Cuando el usuario mira un botón en particular, la señal capturada en la corteza visual contiene esta frecuencia, por lo que se puede hacer la estimación de a qué botón está mirando el usuario. Otro método de uso frecuente es el P300. Se basa en el hecho de que un pico positivo aparece en el EEG aproximadamente 300 ms después de la presentación de un estímulo raro. Aunque aquí tenemos un estímulo externo, se considera que el P300 es un BCI endógeno, ya que la aparición del pico no se vincula con los atributos físicos de un estímulo, sino que refleja procesos involucrados en la evaluación o categorización del estímulo.

3.5 Control del sistema vía motor de recomendaciones

En la introducción se mencionó, que un motor de recomendaciones puede predecir las preferencias del usuario y así ahorrarle tiempo. El sistema puede hacer esto en función de lo que los usuarios han demostrado anteriormente de hacia que se sienten atraídos (lo que más les gusta a los usuarios), lo que les ha gustado a los usuarios similares a ellos (los usuarios se dividen en varios segmentos según sus preferencias) o una combinación de las primeras opciones. Al recopilar todos estos datos, podemos tener información relevante para su interés seleccionado.



El sistema de recomendaciones se vuelve más preciso a medida que aumenta la cantidad de elementos calificados por los usuarios. Además, cuanto más preciso sea el motor, más usuarios se animarán a usarlo. Como resultado, parece que construir un sistema válido es una tarea decisiva.



Un sistema de recomendación es un esquema que predice la calificación de un elemento por un usuario y, al hacerlo, puede elegir la mejor opción para él o ella.

Hay dos principales enfoques: *filtrado colaborativo* (**CF**, del inglés collaborative filtering) y *basado en contenido* (**CB**, del inglés content-based). En la primera, CF, la recomendación se basa en calificaciones de usuarios similares a la actual. En la última, CB, las recomendaciones se basan en las descripciones de los elementos que el usuario calificó previamente.

Ambos enfoques tienen sus ventajas y sus inconvenientes.



CF sufre de un problema de arranque en frío. Es más evidente cuando se lanza un nuevo elemento, ningún usuario lo habrá calificado previamente, por lo que el sistema no lo recomendará, o cuando haya un nuevo usuario que no haya calificado ningún elemento, el sistema no se puede compararlo con ningún otro usuario. Además, encontrar un conjunto de usuarios afines no siempre es una tarea fácil, ya que la probabilidad de que varios usuarios califiquen los mismos elementos es baja. CB carece de novedad y necesita guardar la descripción del contenido para cada elemento.

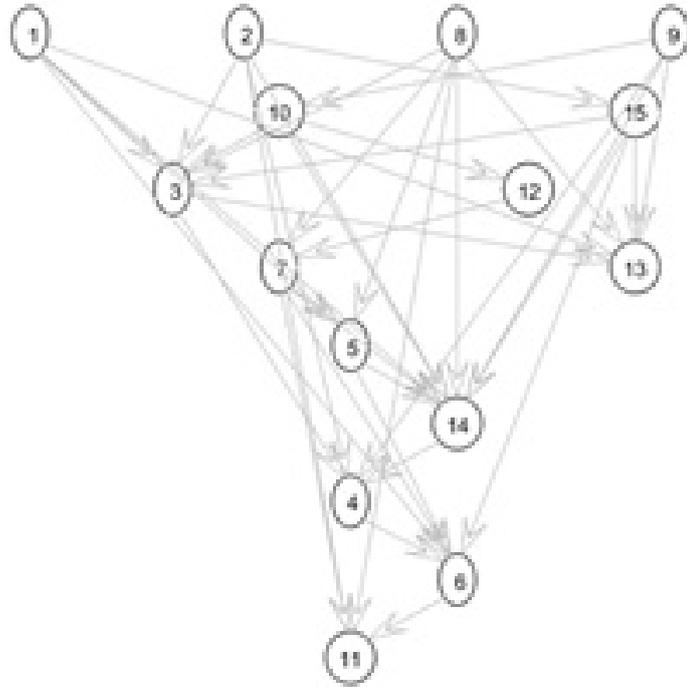


Pero las recomendaciones descubiertas por la CF son fortuitas. Sin embargo, para CB, el problema del arranque en frío se resuelve haciendo que un nuevo usuario complete una breve encuesta.

Se han hecho esfuerzos para combinar ambos enfoques de varias maneras con el fin de minimizar los inconvenientes. Estas técnicas incluyen cambiar entre los dos enfoques, ponderar la salida de ambos esquemas, usarlos en cascada, etc.



Las redes bayesianas se vuelven muy útiles al pensar en modelar nuestro sistema de Recomendaciones. Una red bayesiana es un gráfico acíclico dirigido donde los nodos representan un conjunto de variables aleatorias y los arcos representan dependencias directas entre las variables. La fuerza de la relación entre ellos se cuantifica por la distribución de probabilidad condicional asociada con cada nodo.



Ejemplo de una red Bayesiana

4 4 Control del sistema en aplicaciones

4.1 Comandos de voz para dispositivos móviles

Los dispositivos móviles son actualmente el tipo de dispositivo más utilizado de la tecnología informática de consumo. Abarcan una amplia gama de usos posibles, desde realizar una simple llamada a realizar tareas de trabajo hasta crear y consumir contenido multimedia. Aunque la mayoría de los casos de uso se basan en una pantalla táctil como método principal de entrada de comandos, hay situaciones en las que es necesario el controlarlos con las manos libres.

Conducción

En un vehículo, el conductor está obligado a prestar atención visual a la situación en la carretera. Esto significa que no puede controlar el teléfono con las manos, ya que están en el volante o en otro control del automóvil. La utilización de comandos de voz es necesaria en varias áreas:

- Solicitar navegación al destino deseado
- Realizar una llamada telefónica
- Controlar la reproducción de música
- Otro control de aplicación específico



Los actores más importantes en este campo incluyen Apple, Google, Microsoft o Amazon, cuyos asistentes de voz pueden acceder a la información ubicada en la memoria del dispositivo, así como en Internet. Esto permite realizar una llamada, iniciar una aplicación de navegación con un destino preestablecido, iniciar la reproducción de una canción, pausarla, pasar a la canción siguiente o la anterior y muchas otras funciones básicas. La interacción con cualquiera de los asistentes es en su mayoría fluida y exitosa en la ejecución del comando deseado, y cada uno de ellos contempla las entradas en varias formas de expresión.

Obtención de información

Los asistentes de voz pueden realizar búsquedas en Internet e informar de los resultados de forma lógica, rápida y completa a través del sintetizador de voz. Los comandos se pueden solicitar de una manera compleja, similar a una oración, y la red neuronal deducirá la intención.

Automatización del hogar

El dispositivo móvil también puede servir como un centro de control domótico. Mientras están conectados a la red doméstica, los asistentes dirigidos por voz pueden recibir comandos para controlar varios elementos de la casa, dado que estos elementos están conectados en línea. Las tareas comunes pueden incluir encender o apagar las luces (específicas), cerrar las cortinas, o hasta configurar un ambiente específico con la combinación de varios dispositivos inteligentes.

Necesidades especiales

Las personas con necesidades especiales a menudo requieren interfaces ajustadas a su condición física. Cuando el teclado o los dispositivos táctiles son difíciles de usar, debido a impedimentos visuales o daños en las extremidades, el entorno controlado por voz es una ayuda esencial para operar cualquier sistema multimedia o de información. Dicho sistema generalmente consiste en el sistema de reconocimiento del comando de voz y la retroalimentación del mismo.

4.2 Gestos para sistemas modernos de control de TV

Uno de los mayores inconvenientes de un uso más amplio de las interfaces de usuario naturales es su falta de usabilidad y diseño centrado en el ser humano. Si bien otras modalidades (es decir, la navegación por comandos de voz) parecen adaptarse bastante rápidamente, el reconocimiento de gestos todavía no puede ofrecer una experiencia verdaderamente natural, especialmente en dispositivos sin contacto. Hay varios factores que determinan la intuición y la naturalidad del reconocimiento de gestos. En primer lugar, son las limitaciones de hardware las que limitan los algoritmos de los sensores para reconocer detalles más específicos en los gestos realizados. Esto hace que los gestos se reconozcan incorrectamente y obliga a los usuarios a realizar gestos que requieren mucho esfuerzo y carecen de comodidad. En segundo lugar, los conjuntos de gestos propuestos actualmente en sistemas sin contacto no son intrínsecamente intuitivos. Los diseñadores de sistemas tienden a superar las limitaciones de los sensores al introducir gestos que son fácilmente reconocibles, pero a menudo están lejos de ser simples [2].

Apple TV

Apple TV usa su control remoto para captar gestos. La superficie táctil del control remoto detecta una variedad de gestos intuitivos con un solo dedo. Hay tres tipos de gestos.



Deslizar. Mueve el foco hacia arriba, hacia abajo, hacia la izquierda o hacia la derecha entre los elementos. Al deslizar, el usuario puede desplazarse sin esfuerzo a través de grandes volúmenes de contenido con un movimiento que comienza rápido y luego se ralentiza, según la fuerza del deslizamiento.

Clic. Activa un control o selecciona un elemento. Hacer clic es la forma principal de desencadenar acciones. Hacer clic y mantener a veces se usa para desencadenar acciones específicas del contexto. Por ejemplo, hacer clic y mantener presionado un elemento de la interfaz puede hacer entrar en un modo de edición.

Toque. Navega a través de una colección de artículos uno por uno. En aplicaciones con interfaces estándar basadas en UIKit, al tocar diferentes regiones se navega hacia una dirección. Por ejemplo, al tocar la parte superior de la superficie táctil, se navega hacia arriba. Algunas aplicaciones utilizan gestos de toque para mostrar controles ocultos.

Diferenciar entre hacer clic y tocar, y evitar desencadenar acciones con toques involuntarios. Hacer clic es una acción que requiere mucha intención, y generalmente es muy adecuado para presionar un botón, confirmar una selección e iniciar una acción durante el juego. Los gestos de toque son adecuados para la navegación o para mostrar información adicional, pero tenga en cuenta que el usuario, naturalmente, puede colocar un pulgar en el control remoto, levantarlo, moverlo o dárselo a alguien.



Controlador de Apple TV

SingleCue

El Singlecue original se lanzó a fines de 2014 y ofreció una visión temprana de lo que podría ser posible con el control por gestos. Con el tamaño de una Xbox Kinect, ese dispositivo funcionaba a través de infrarrojos y te permitía encender un dispositivo con una onda, silenciar el volumen al ponerte el dedo en los labios o cambiar entre dispositivos con una variedad de otros gestos.

El Singlecue de segunda generación se basa en el de primera agregando nuevos gestos, como un gesto con la mano, un pellizco con el dedo y un clic con la palma para hacer que el dispositivo sea aún más útil para el usuario final.



Por ejemplo, ahora se puede controlar la reproducción y la pausa del vídeo abriendo y cerrando la mano, mientras que el volumen se puede controlar moviendo un dedo contraído de izquierda a derecha. Estos gestos funcionarían en cualquier momento, lo que significa que el usuario no necesariamente tendría que estar en un menú específico para acceder a esa funcionalidad.



Sensor SingleCue

Samsung TV / LG TV

Samsung TV utiliza un control simple por gestos para acceder a sus películas, deportes, aplicaciones favoritas y a otros contenidos inteligentes en Samsung Smart TV. Samsung usa una cámara simple para monitorizar el entorno frente a ella.



El usuario puede olvidar el control remoto y usar sus manos para controlar las funciones de TV deslizando para navegar y agarrando para seleccionar. Es tan inteligente como fácil. Se puede utilizar el control de movimiento para cambiar el canal, ajustar el volumen, mover el puntero y controlar otras funciones del televisor. Los gestos compatibles son deslizar, hacer zoom, agarrar.

Este conjunto de gestos permite el control básico de una Smart TV. El movimiento puede estar limitado por:

- condiciones de luz oscura o demasiado brillante,
- estar demasiado cerca o demasiado lejos de la cámara,
- los dedos no pueden ser detectados en el caso de que se esté usando guantes o un vendaje,
- tener la mano frente a la cara durante el reconocimiento de movimiento,
- haber luz solar directa,
- usar otros dedos en lugar del dedo índice.



Control por gestos de Samsung TV

4.3 Gestos para teléfonos inteligentes (y otras aplicaciones)

Muchas aplicaciones en teléfonos móviles han diseñado e implementado el uso de gestos intuitivos que permitirían a los usuarios imaginar qué movimiento deberían realizar para ejecutar un comando específico. Los gestos también permiten a los diseñadores desarrollar interfaces agradables al permitir más posibilidades para las ideas profesionales. Las nuevas interfaces están ahora generalmente diseñadas sin botones para hacer clic y permiten posibilidades para ideas profesionales. Los botones no pueden desaparecer de la aplicación móvil para siempre, ya que juegan un papel crucial en la petición de acciones. Sin embargo, en el caso de que los gestos sean más naturales e intuitivos, y simplifiquen la interacción del usuario, deben implementarse [5].



Un ejemplo de aplicación móvil, controlada con gestos, puede ser Google Maps o cualquier sistema de navegación utilizado en el teléfono móvil. Una aplicación como Google Maps brinda a los usuarios la oportunidad de aplicar varios gestos para controlar determinadas funciones suyas. Por ejemplo, para acercar o alejar el mapa en la pantalla, puede usarse el dedo moviéndolo hacia arriba y hacia abajo respectivamente.

Otra aplicación es Clear, una tarea de iOS que administra la aplicación móvil. El hecho sorprendente de esta aplicación es que no tiene botones, por lo que se basa completamente en el control de gestos. Utiliza toques y desliza para agregar y eliminar tareas de una lista de tareas pendientes.

Definitivamente hay más aplicaciones basadas en la navegación y control por gestos, pero enumeramos solo las más utilizadas.



Google maps con navegación por gestos

Otras aplicaciones que utilizan el control gestual

Muchas aplicaciones se desarrollan para aplicaciones de entrenamiento profesional (por ejemplo, golf, béisbol). El usuario no necesita ningún hardware adicional como teclados y joysticks. Las aplicaciones pueden proporcionar capacidades de captura tridimensional del movimiento del cuerpo y la mano en tiempo real. Para este propósito, se utiliza Kinect. Kinect usa una cámara de profundidad para control de movimiento. Mientras que Kinect se centra principalmente en capturar la pose corporal, Leap Motion desarrolló un dispositivo de captura de gestos de corto alcance con una cámara infrarroja estéreo. Leap Motion puede seguir los gestos finos de las dos manos a una alta velocidad de fotogramas por segundo. Permite aplicaciones como dibujar y manipular objetos pequeños en un espacio virtual. Algunos proveedores de PC se asociaron con Leap Motion para proporcionar una interfaz de usuario natural (NUI, del inglés Natural User Interface) en aplicaciones de escritorio como el *Diseño Asistido por Computador (CAD, del inglés Computer Aided Design)* [3].

Varios proveedores de software también proporcionan un *Kit de Desarrollo Software (SDK, del inglés Software Development Kit)* o software intermedio (middleware en inglés) para que los desarrolladores de aplicaciones integren fácilmente el reconocimiento de gestos y posturas en sus aplicaciones (por ejemplo, Gestoos, eyeSight).

Para eliminar las distracciones del conductor y aumentar la seguridad del tráfico, los fabricantes de automóviles están creando una interfaz sin contacto basada en los gestos de la mano (por ejemplo, el sistema de control de gestos basado en cámaras de BMW). Esta interfaz reduce la necesidad de que los conductores lleguen al panel de control del tablero. Es una forma más natural de controlar el sistema de

información y entretenimiento y ayuda a mantener la atención del conductor en la carretera.

Los gestos con las manos son la forma viable de guiar también las operaciones de aviones no tripulados (drones en inglés) en exteriores, de tal manera que los aviones no tripulados pueden volar de forma autónoma y ser guiados sin utilizar el control remoto (por ejemplo, volver a llamar al avión no tripulado agitando las manos). Un ejemplo puede ser un nuevo dron de DJI llamado Spark. Spark es un dron controlado por gestos de la mano, donde todo lo que necesitas son tus manos para comandarlo. Se puede ordenar que se aleje, que nos tome una foto o que explore libremente los cielos de la forma que se desee, todo con gestos. Spark puede detectar objetos delante de él a una distancia de hasta 5 m, para evitar automáticamente cualquier colisión desagradable [4].



Spark

4.4 Seguimiento ocular (control del ratón del PC, navegación entre opciones en pantalla)

La opción de seguimiento ocular para el control del ratón ha surgido en los últimos años como una opción de accesibilidad para controlar el PC. Ya está integrado en algunos sistemas operativos, por ejemplo en Windows 10 de Microsoft.

Windows 10 es compatible con hardware específico de seguimiento ocular. Algunas funcionalidades clave se pueden controlar a través del movimiento del ojo. Estas funcionalidades se centran en acceder a las aplicaciones, introducir información y comunicarse. El elemento básico de la *interfaz de usuario (UI*, del inglés User Interface) de control mediante los ojos de Windows 10 es el launchpad, que permite a los usuarios acceder al ratón, al teclado y a las funciones de texto a voz simplemente mirando los iconos, así como mover el launchpad hacia arriba o hacia abajo por la pantalla. Interactuar con la interfaz de usuario de Control Ocular es simple pero requiere algo de práctica: basta con mirar la pantalla y enfocar, o "detener" como Microsoft lo llama, en un botón u otro elemento hasta que se active. El sistema proporciona comentarios que permiten saber lo que se está haciendo.

Como reemplazo del ratón para un usuario genérico, la precisión de estos sistemas no es suficiente, pero en los casos de usuarios con necesidades especiales, para los cuales una pantalla magnificada y una operación más lenta no son un problema, esta alternativa puede ser una gran elección.

Cuando se usa para el control en juegos que no requieren el control fino de un tirador en primera persona, el jugador puede lograr una experiencia muy inmersiva, navegando simplemente mirando hacia dónde quiere ir el jugador. La cámara puede permitir que el jugador mire donde quiera o se enfoque en un personaje específico. En tales situaciones, el jugador podría sentirse como si estuviera dentro del juego.

Hay muchas más áreas de aplicación, no solo control de ordenadores. Por ejemplo actualmente se está desarrollando una tecnología de seguimiento ocular llamada "cabina de mando portátil", para los futuros aviones de combate que proporcionará a los pilotos de aviones de combate una pantalla virtual proyectada a través del casco. Permitirá a los pilotos acceder, evaluar y actuar rápidamente en base a información crítica, proporcionando un control sencillo en la cabina de la aeronave. Por ejemplo con solo mirar algo se podría resaltar y luego haciendo un gesto para "presionar" se podría presionar.

También hay algunos programas que siguen la posición de la cabeza usando la cámara del portátil y la utilizan para el control del ratón (por ejemplo, iTracker).

4.5 BCI (control de sillas de ruedas, navegación en juegos)

Hay más ejemplos donde se utiliza el BCI. Este es en muchos casos la solución entusiasta, principalmente para controlar algunos drones, robots y modelos de coches. En otros casos, es un dispositivo de asistencia para personas con discapacidades como en el control de sillas de ruedas o de la máquina de deletreo. En el siguiente texto se dan algunos casos representativos.

- Deletreador P300. En el deletreador P300 se usa extensivamente el método P300 (pico positivo que aparece en el EEG aproximadamente 300 ms después de la presentación de un estímulo raro). Está formado por un teclado virtual en pantalla. El usuario mira la letra deseada. Los grupos de teclas parpadean de forma secuencial hasta que se reconoce la letra. En algunas soluciones clásicas de P300, los grupos están formados por filas y columnas clave, en las soluciones más nuevas se trata de grupos de letras pseudoaleatorias diseñados para minimizar el parpadeo consecutivo de los mismos caracteres y el parpadeo simultáneo de los caracteres vecinos. Con estos sistemas, se pueden escribir alrededor de 3.5 letras / minuto [8] con un 95% de letras correctas.
- Control de silla de ruedas. Para el control de sillas de ruedas, se usan principalmente sistemas BCI basados en P300, SSVEP y sensoriomotores [9].
- Control del hogar inteligente. Hay más estudios donde se analiza el BCI integrado en algunos sistemas de automatización de hogares inteligentes. Por ejemplo en un estudio se analizaron los controles de la iluminación, el televisor, la cafetera y las persianas del hogar, y se encontró un 80% de precisión en la tarea. Los métodos más utilizados en los estudios son P300 y SSVEP [10].
- Control de prototipos. Hay una gran cantidad de informes y estudios en los que se utilizó el BCI para el control de prototipos (drones, automóviles, robots, ...). Ninguna solución comercial está disponible hasta hoy, todavía es un área de investigación activa. Se dispone de una amplia gama de soluciones difíciles, pero la facilidad de uso es discutible. En los últimos años se organizaron también más carreras, por ejemplo carreras de coches, carreras de drones [11] utilizando el control basado en BCI.
- Lectura de emociones. Aunque los sistemas BCI pueden hasta cierto punto "leer emociones". De vez en cuando aparecen artículos en los periódicos sobre la disponibilidad y usabilidad de la tecnología, incluso en grandes implementaciones [12], pero luego aparecen otros, llenos de dudas [13]. BCI se puede usar para determinar si alguien está despierto o dormido, pero para estados emocionales complejos como la depresión y la ansiedad, aún no se ha entendido bien qué patrones de actividad cerebral coinciden con qué etapas emocionales [14].
- Control de juegos. Hay muchos proyectos que intentaron introducir la interfaz de control de juegos basada en BCI. Entre las más exitosas están las soluciones basadas en SSVEP [14].

En la mayoría de los casos, es difícil juzgar si el sistema para el caso de uso particular es realmente utilizable. Debido a los principales problemas relacionados con la comodidad, la velocidad y la confiabilidad, definitivamente no son una alternativa competitiva para usuarios sin problemas. Sin embargo, esto puede cambiar muy pronto cuando, por ejemplo, si la etapa de adquisición de señal pudiera realizar muchos progresos.

5 5 ¿Conclusión o quo vadis, control del sistema?

La mayoría de las herramientas, que las personas han desarrollado en la historia, constan de dos partes principales, la parte funcional de la herramienta y, en general, un interfaz o facilitador del uso. Un ejemplo típico es un cuchillo que tiene una cuchilla y un agarre. Cuanto más compleja es la herramienta desde el punto de vista de la funcionalidad, más sofisticada es la interfaz interactiva para el control del sistema.

Las aplicaciones informáticas representan un tipo de herramienta muy específico con una funcionalidad generalmente muy complicada. Este hecho resulta en la necesidad de una interfaz específica para el control del sistema. El código hexadecimal se puede considerar como una de las primeras interfaces seguidas por varios lenguajes de programación para el manejo de computadoras. El desarrollo de sistemas operativos modernos es un esfuerzo notable para que trabajar con computadoras sea lo más cómodo posible. La introducción de la filosofía de Windows y el ratón fue un primer paso para simplificar para los humanos la comunicación más que simplificar las computadoras.

La investigación en el campo de la comunicación natural humana ha conducido a sistemas biométricos que tienen un origen en la comunicación humana en la vida real. Las modalidades típicas en la dirección humano-computadora son el control por voz, el control por gestos, la identificación del hablante, el reconocimiento facial, el reconocimiento del movimiento corporal, etc. En la dirección opuesta, modalidades como el sistema de advertencia / acuerdo / desacuerdo, síntesis de voz para retroalimentación de voz, avatar gráfico en forma de varias criaturas o avatares humanoides para la interacción visual y en los últimos años, la realidad virtual (VR, del inglés Virtual Reality) y la realidad aumentada (AR, del inglés Augmented Reality) se utilizan para proporcionar información al usuario. Se ha realizado mucho esfuerzo para desarrollar y utilizar tales modalidades, que siguen la expectativa humana sobre cómo entendemos la comunicación natural. Todos los tipos de interfaces que son naturales para las computadoras no son naturales para los humanos y viceversa.

Conseguir la facilidad y calidad de uso de la interfaz hombre-computadora (HCI, del inglés Human-Computer-Interface) es a menudo un problema más complejo que implementar el método en sí en la aplicación. Lograr un HCI fácil de usar y natural, donde el usuario pueda comunicarse de una manera natural no es una etapa final del desarrollo de HCI. El siguiente paso es un sistema, que observe al usuario y su comportamiento como una actividad a corto plazo o mejor a largo plazo. El llamado motor de recomendación será capaz de predecir no solo el comportamiento futuro del usuario, sino también sus deseos y necesidades reales y activar las funciones / comandos relevantes para satisfacer al usuario. La computadora poseerá las funciones de la inteligencia artificial y se convertirá no solo en una máquina sino en un socio.