

Consulte los debates, las estadísticas y los perfiles de los autores de esta publicación en: <https://www.researchgate.net/publication/330617900>

BBEEP: Un sistema sónico de prevención de colisiones para viajeros ciegos y peatones cercanos

Ponencia de la conferencia - mayo de 2019

DOI: 10.1145/3290605.3300282

CITACIONES

106

LEA

975

7 autores, entre ellos:



[Seita Kayukawa](#)

IBM Investigación

18 PUBLICACIONES 238 CITAS

VER PERFIL



[Keita Higuchi](#)

Universidad de Tokio

28 PUBLICACIONES 756 CITAS

VER PERFIL



[João Guerreiro](#)

Universidad de Lisboa

72 PUBLICACIONES 1.217 CITAS

VER PERFIL



[Shigeo Morishima](#)

Universidad de Waseda

368 PUBLICACIONES 3.949 CITAS

VER PERFIL

BBEEP: Un sistema sónico de prevención de colisiones para viajeros ciegos y peatones cercanos

Seita Kayukawa
Universidad de Waseda

Universidad Carnegie Mellon
k940805k@ruri.waseda.jp

Keita Higuchi
Universidad de Tokio

khiguchi@iis.u-tokyo.ac.jp

João Guerreiro
Universidad Carnegie Mellon

jpguerreiro@cmu.edu

Shigeo Morishima
Instituto de Investigación Waseda para Ciencia e ingeniería

shigeo@waseda.jp

Yoichi Sato
Universidad de Tokio
ysato@iis.u-tokyo.ac.jp

Kris Kitani
Universidad Carnegie Mellon
kkitani@cs.cmu.edu

Chieko Asakawa
Investigación de IBM
Universidad Carnegie Mellon
chiekoa@cs.cmu.edu

RESUMEN

Presentamos un sistema de maleta de asistencia, BBEEP, para ayudar a las personas ciegas a caminar por entornos abarrotados. BBEEP utiliza notificaciones sonoras preventivas para ayudar a despejar el camino alertando tanto al usuario como a los peatones cercanos sobre el riesgo potencial de colisión. BBEEP activa las notificaciones rastreando a los peatones, prediciendo su posición futura en tiempo real y emitiendo notificaciones sonoras sólo cuando prevé una futura colisión. Investigamos cómo afectan los distintos tipos y tiempos de sonido al comportamiento de los peatones cercanos. En nuestros experimentos, descubrimos que el momento de emisión del sonido tiene un impacto significativo en las trayectorias de los peatones cercanos en comparación con los distintos tipos de sonido. Basándonos en estos resultados, realizamos un estudio con usuarios reales en un aeropuerto internacional, en el que participantes ciegos navegaban con la maleta en zonas abarrotadas de gente. Observamos que el sistema propuesto reduce significativamente el número de colisiones inminentes.

CONCEPTOS DE CSS

• **Informática centrada en el ser humano** → **Tecnologías de accesibilidad**; - **Temas sociales y profesionales** → *Personas con dis- capacidad.*

PALABRAS CLAVE

Discapacidad visual; evitación de obstáculos; predicción de colisiones; detección de peatones; navegación a ciegas; despeje de trayectorias.

Se autoriza la realización de copias digitales o impresas de la totalidad o parte de esta obra para uso personal o en el aula sin coste alguno, siempre que las copias no se realicen o distribuyan con fines lucrativos o comerciales y que las copias lleven este aviso y la cita completa en la primera página. Deben respetarse los derechos de autor de los componentes de esta obra que no pertenezcan a ACM. Se permite hacer resúmenes con los créditos correspondientes. Cualquier otra copia, republicación, publicación en servidores o redistribución a listas requiere

un permiso específico previo y/o el pago de una tasa. Solicite permiso a permissions@acm.org.

CHI 2019, del 4 al 9 de mayo de 2019, Glasgow, Escocia Reino Unido.

2019 Association for Computing Machinery. ACM ISBN

978-1-4503-5970-2/19/05...\$15.00

<https://doi.org/10.1145/3290605.3300282>



Figura 1: BBEEP es un sistema de asistencia para maletas que utiliza la retroalimentación sónica para alertar tanto al usuario ciego como a los peatones videntes cercanos sobre posibles riesgos de colisión.

Formato de referencia ACM:

Seita Kayukawa, Keita Higuchi, João Guerreiro, Shigeo Morishima, Yoichi Sato, Kris Kitani y Chieko Asakawa. 2019. BBEEP: Un sistema sónico de evitación de colisiones para viajeros ciegos y peatones cercanos. En *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, 4-9 de mayo de 2019, Glasgow, Escocia Reino Unido. ACM, Nueva York, NY, Estados Unidos, 12 páginas. <https://doi.org/10.1145/3290605.3300282>

1 INTRODUCCIÓN

Las personas ciegas se enfrentan a importantes retos a la hora de desplazarse por espacios públicos debido a la falta de percepción visual. Recientemente se han realizado búsquedas mediante visión por computador con el objetivo de ayudar a los usuarios ciegos a orientarse y moverse para evitar posibles obstáculos o peligros [7, 16, 22, 23, 29, 37, 46, 49, 50]. Aunque estos sistemas suelen ser capaces de detectar obstáculos estáticos, la detección

y la evitación de colisiones con elementos dinámicos, en particular peatones, sigue estando relativamente inexplorada en la literatura. Dejando a un lado los problemas técnicos, una posible razón de la falta de trabajos sobre elementos dinámicos es la suposición de que los peatones videntes son conscientes de la presencia de personas ciegas y, por tanto, siempre les despejarán el camino. Sin embargo, esto no siempre es así, ya que las personas videntes pueden estar mirando su teléfono inteligente, hablando con otras personas o mirando en otra dirección (un tablero o la televisión). En tales situaciones, los invidentes corren un riesgo considerable de colisión con otros peatones.

Presentamos un sistema de asistencia para maletas, BBleep, que utiliza un sistema de aviso sonoro de colisión para alertar tanto al usuario ciego como a los peatones videntes cercanos sobre posibles riesgos de colisión (figura 1). Este enfoque amplía los sistemas de advertencia sónica habituales que se utilizan para despejar el camino a los vehículos en movimiento, como los carritos de los aeropuertos que circulan por terminales abarrotadas o los grandes camiones que circulan marcha atrás. Este trabajo aprovecha el simple hecho de que las personas videntes pueden apartarse rápidamente del camino de una persona ciega que camina, si reciben información adecuada sobre la presencia de una persona ciega. Sin embargo, nuestro trabajo pretende ir más allá del paradigma de reproducir constantemente un sonido para transmitir la presencia del usuario, ya que la emisión constante de fuertes sonidos de alarma puede ser socialmente perturbadora y hacer que el usuario ciego se sienta demasiado cohibido. En su lugar, presentamos un sistema de aviso sonoro adaptativo que sólo emite sonidos cuando es necesario. Más concretamente, BBleep está diseñado para tener en cuenta el movimiento de los peatones cercanos, predecir futuras colisiones y emitir avisos sonoros sólo cuando sea necesario.

Aunque nuestro objetivo explícito es la navegación de invidentes en aeropuertos, creemos que el factor de forma de una maleta de viaje también es apropiado y aporta varias ventajas en otros entornos abarrotados del mundo real, como estaciones de tren o centros comerciales. Para el usuario ciego, una maleta puede actuar a menudo como un mecanismo de detección ampliado para identificar cambios en la textura del suelo o como una forma de protección contra colisiones en aglomeraciones muy densas. Incluso sin sensores inteligentes, una maleta puede utilizarse como dispositivo de asistencia. En muchos entornos cosmopolitas, una maleta es un objeto común y no llama innecesariamente la atención del usuario. Como sistema de detección robótico, también ofrece un lugar práctico para almacenar y conectar sensores, energía y recursos informáticos.

BBleep utiliza una cámara RGBD para detectar, seguir y predecir el movimiento de los peatones cercanos. La imagen RGB se utiliza para detectar personas mediante una red neuronal convolucional y el canal de profundidad se utiliza para estimar la distancia a los peatones. La posición media estimada se utiliza para calcular la velocidad de los peatones y la extrapolación lineal para predecir su

trayectoria. En función de la proximidad de la trayectoria prevista con el usuario, BBleep emite un sonido adecuado.

Para investigar cómo transmitir eficazmente la retroalimentación sonora, realizamos un estudio observacional en el que el sistema en forma de maleta emite sonidos de diferentes tipos y tiempos.

Los resultados sugieren que la emisión de sonidos es un método eficaz para cambiar la dirección de marcha de los peatones alejándolos del andén, y que su sincronización tiene más impacto que los tipos de sonido. Basándonos en estos resultados, diseñamos la interfaz sonora de BBeep, que utiliza tres etapas de emisiones sonoras para notificar los posibles riesgos de colisión con los peatones.

Para evaluar la eficacia de BBeep en la prevención de colisiones con peatones, realizamos un estudio en el que seis usuarios ciegos caminaban con la maleta por zonas concurridas de un aeropuerto. Observamos que BBeep reducía el número de situaciones de riesgo de colisión inminente, en comparación con la única notificación al usuario ciego. Los comentarios de los participantes respaldaron nuestra hipótesis de que BBeep es útil para evitar colisiones en espacios públicos abarrotados. Basándonos en nuestros hallazgos, discutimos los requisitos futuros hacia una solución más flexible que sea capaz de adaptarse a diferentes escenarios y usuarios.

2 TRABAJOS RELACIONADOS

Sistemas de navegación a ciegas

Investigaciones anteriores han propuesto varios sistemas de navegación para ciegos [1, 2, 10, 13, 18, 31, 35, 36, 39]. La mayoría de los sistemas guían a las personas ciegas mediante navegación giro a giro, con tecnologías como GPS [36, 39], etiquetas RFID [2, 10, 15] y balizas Bluetooth de baja energía [1, 11, 26]. Sin embargo, la mayoría de los sistemas no tienen en cuenta los entornos dinámicos y, por tanto, no son conscientes de los obstáculos que no se encuentran en el entorno, como escritorios, sillas y peatones. Para superar esta limitación, es importante ayudar a los usuarios ciegos a evitar colisiones con tales elementos.

Sistemas de apoyo para evitar obstáculos

Además de los perros guía, los bastones blancos [5, 47] son la herramienta más común para que los invidentes encuentren obstáculos y eviten colisiones. Aunque es muy eficaz, un usuario sólo puede detectar un obstáculo después de golpearlo físicamente con el bastón. Esto no es deseable, sobre todo cuando el obstáculo es un peatón. Los investigadores han desarrollado tecnologías de apoyo que permiten a los usuarios ciegos detectar obstáculos sin contacto [4, 30, 43]. Los sistemas que detectan obstáculos y proporcionan información sobre ellos (por *ejemplo*, distancia [22-24, 28, 30, 37], forma [7, 23, 29] o categoría [22, 37, 49]) a los usuarios suelen utilizar láser [28], ultrasonidos [24, 44], altavoces y micrófonos del teléfono [45] o micrófonos [46]. Sin detección de profundidad [7, 16, 21-23, 37, 49]. Sin

embargo, sigue siendo difícil para los invidentes detectar y evitar obstáculos, sobre todo en entornos muy densos. En estos casos, es importante proporcionar un camino seguro a los usuarios ciegos.

Algunos sistemas guían a los usuarios para sortear obstáculos mediante retroalimentación sonora [29, 50] y/o háptica procedente de robots terrestres [46] o aéreos [3]. Los usuarios ciegos pueden seguir la retroalimentación del sistema para evitar obstáculos estáticos como sillas, escritorios y paredes. Aunque son aplicables a objetos inanimados, algunos estudios suponen que los peatones, como obstáculos dinámicos, pueden evitar a un usuario ciego,

y, por tanto, no se centraron en ayudar a los peatones a evitar colisiones [29, 46]. Sin embargo, en espacios públicos (por *ejemplo*, aeropuertos), los peatones pueden no ser conscientes de la presencia de usuarios ciegos mientras utilizan dispositivos móviles o hablan con otras personas. Por lo tanto, estamos interesados en investigar técnicas de notificación sonora para que los peatones sean conscientes de la presencia de un viajero ciego. Esperamos que los peatones despejen el camino para el usuario ciego.

Alerta sonora para notificaciones urgentes

Los pitidos se han utilizado para avisar de situaciones urgentes, como en las unidades de cuidados intensivos de los hospitales [34], las centrales nucleares [33] y la aviación [6]. Las notificaciones sonoras también pueden alertar a los conductores de una colisión inminente o ayudarles en la navegación [32]. La relación entre la percepción del usuario y los distintos tipos de sonidos de alerta desempeña un papel fundamental en su usabilidad. Varios trabajos han descubierto que los parámetros auditivos de los pitidos (por *ejemplo*, la frecuencia fundamental, la frecuencia del pulso y la intensidad) afectan a los niveles de urgencia percibida [12, 20, 32], mientras que otros han observado un equilibrio entre la urgencia percibida y los niveles de molestia de los sonidos de alerta [12, 19, 32].

Como se ha descrito anteriormente, la emisión de pitidos es un método habitual para avisar a los usuarios de situaciones urgentes. En este artículo, utilizamos este tipo de sonido para que los peatones se percaten de la presencia de un usuario ciego. Investigamos qué tipos de pitidos son eficaces para evitar colisiones y diseñamos una política de emisión de sonidos para nuestro sistema prototipo.

3 BBEEP: DISEÑO DE UN SISTEMA DE DESPEJE DE SENDEROS

Nuestro principal objetivo es facilitar la movilidad de las personas ciegas en entornos abarrotados. Creemos que las colisiones con peatones pueden evitarse si tanto el usuario ciego como los peatones videntes son conscientes del riesgo de colisión. Con este fin, hemos desarrollado BBEEP, un sistema sónico para evitar colisiones que pretende despejar el camino a los usuarios ciegos.

Limitaciones de notificar sólo al usuario ciego

Las investigaciones previas sobre detección de obstáculos para invidentes se centran en notificar al usuario por sí solo la presencia de obstáculos, incitándole a cambiar su orientación [29, 46, 50]. Este enfoque aumenta el conocimiento del entorno por parte del usuario, pero también presenta importantes limitaciones en este contexto. En primer lugar, cambiar activamente la dirección de la marcha de los usuarios puede ser inseguro (llevando al usuario por un camino diferente/desconocido); en segundo lugar, un grupo de peatones puede bloquear toda la ruta de un usuario ciego; y, por último, notificar a los usuarios sobre todos los peatones en entornos abarrotados puede requerir una retroalimentación compleja, que puede resultar

cognitivamente exigente para los usuarios. Además, al centrarse en los obstáculos en general, estos enfoques no aprovechan la capacidad de los peatones videntes para cooperar en la prevención de colisiones. Argumentamos que transmitir información sólo al usuario puede no ser eficaz para evitar colisiones con otros peatones, en particular en entornos abarrotados.

Notificaciones sonoras para usuarios y peatones

BBeep utiliza notificaciones sonoras debido a la capacidad del sonido para atraer la atención de las personas incluso cuando están concentradas en otra cosa. Aunque otras modalidades, como los estímulos visuales (por *ejemplo*, Vection Field [17]), también pueden influir en la dirección de la marcha de los peatones, pueden no ser tan eficaces en varios escenarios. Por ejemplo, la prevalencia de los teléfonos inteligentes reduce significativamente la conciencia del entorno de los peatones videntes, lo que puede provocar colisiones [9]. Además, es posible que (grupos de) personas que hablan o miran en otra dirección no se percaten de la presencia de una persona ciega hasta que su bastón blanco les alcance. Por este motivo, nuestro enfoque se inspira en el uso habitual de pitidos para notificar a los peatones situaciones urgentes que les incitan a despejar el camino. Algunos ejemplos son los carritos en aeropuertos abarrotados o los grandes vehículos de motor que circulan marcha atrás. Sin embargo, este enfoque también plantea importantes retos, ya que la emisión frecuente de sonidos de alarma puede resultar socialmente molesta y hacer que el usuario se sienta incómodo.

Predicción de colisiones para reducir las emisiones sonoras

Para mejorar la aceptación social de las emisiones sonoras, es importante emitir sonidos de alerta sólo cuando sea absolutamente necesario. Además, se necesita una técnica de predicción de colisiones para reducir al máximo el número de emisiones sonoras y mantener su eficacia. Por este motivo, BBeep se basa en el seguimiento de peatones en tiempo real y en la predicción de colisiones para emitir notificaciones sólo cuando existe un riesgo potencial de colisión. Esto es beneficioso para reducir tanto los riesgos de colisión como las perturbaciones sociales en los espacios públicos.

4 BBEEP: APLICACIÓN

Hemos desarrollado un sistema de detección por visión para seguir el movimiento de los peatones y predecir su posición futura con el fin de generar una señal de advertencia sonora que despeje el camino delante de un usuario ciego (Figura 2). Se instala una cámara estereoscópica en una maleta para captar imágenes RGB y recoger datos de profundidad. Una de las ventajas de esta configuración es la posibilidad de captar imágenes sin que el movimiento provoque un desenfoque significativo y de realizar los cálculos necesarios *in situ*. El sistema detecta a los peatones con imágenes RGB y rastrea su posición con datos de profundidad en tiempo real. A partir de estos resultados, el sistema predice la posición futura de

los peatones y determina el riesgo de colisión con un usuario ciego. En caso necesario, el sistema emite una alarma acústica.

Observamos que existen trabajos previos sobre la predicción de la trayectoria de los peatones [25, 27, 48] y el objetivo de este trabajo no es avanzar en el estado de la técnica a este respecto. En cambio, nuestra contribución es el análisis y desarrollo de mecanismos eficaces de retroalimentación sónica basados en dicha información predictiva. Para ello, nuestro principal reto es desarrollar una técnica de predicción en tiempo real con suficiente precisión para la predicción de colisiones.

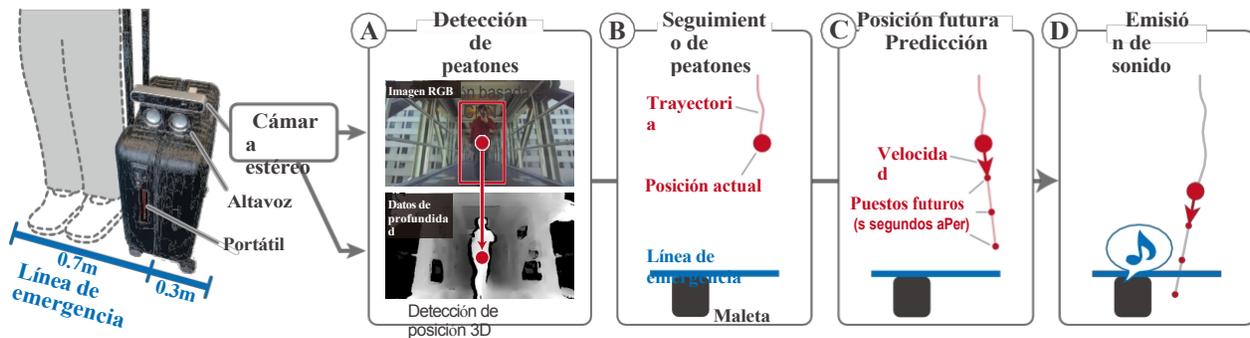


Figura 2: Visión general de BBEEP. La cámara estereoscópica se monta en una maleta y graba imágenes RGB y datos de profundidad. A) El sistema detecta a los peatones utilizando las imágenes RGB y B) rastrea su posición utilizando los datos de profundidad. A continuación, C) predice las posiciones futuras de cada peatón. Por último, D) BBEEP emite una señal acústica si existe riesgo de colisión con el usuario ciego.

Detección de peatones: Requisitos y diseño

Hemos implementado un novedoso sistema que combina la detección de imágenes estereoscópicas y un detector de objetos genérico basado en una CNN (YOLOv2 [40]). Utilizamos una cámara estereoscópica ZED™2K1, ya que tiene un campo de visión horizontal más amplio (90 grados), un mayor rango de profundidad (de 0,3 a 0,5 metros) y una mayor capacidad de detección de objetos.

20 m), y unos FPS más altos (hasta 100 Hz) que Kinect v2.

La cámara estéreo admite una API de odometría 3D que proporciona acceso a los movimientos 3D de la cámara en tiempo real. Utilizamos esta información para eliminar la influencia de la rotación de la maleta. Utilizamos YOLOv2 para detectar peatones utilizando los flujos RGB. El método detecta con solidez a los individuos aunque su cuerpo no esté completamente incluido en las imágenes de la cámara.

Confirmamos que la combinación de la cámara ZED y YOLOv2 puede detectar recuadros delimitadores de peatones a una distancia de 10 m. Utilizamos el área central de los recuadros delimitadores para obtener las posiciones 3D de los peatones detectados.

en el sistema de coordenadas de la cámara.

Para predecir con precisión la posición futura de los peatones es importante actualizar las posiciones a una velocidad alta. Utilizamos un ordenador portátil (CPU Intel Core i7-7700HQ, GPU NVIDIA GeForce GTX 1060) para procesar la detección de objetos a una velocidad de 15 fps, pero esto fue insuficiente para obtener una precisión de predicción satisfactoria. Por lo tanto, utilizamos un cuadro delimitador determinado para obtener las posiciones 3D y lo actualizamos cada vez que disponíamos de un nuevo resultado de detección. De este modo, el sistema rastrea a los peatones a una velocidad superior a 40 fps. Los procesos de detección y seguimiento se ejecutan simultáneamente en subprocesos distintos.

Seguimiento de peatones

El sistema procesa el seguimiento de peatones basándose en los resultados de la detección. Proponemos un algoritmo para rastrear peatones en tiempo real. Utilizamos el siguiente

en el sistema de coordenadas de la cámara basándose en el área central de los recuadros delimitadores.

- (2) El rastreador repite los pasos (3) y (4) para cada posición 3D detectada (la posición actual).
- (3) Si no hay peatones en la lista desde la posición actual a una distancia α , el rastreador añade el punto a la lista como nuevo peatón.
- (4) En caso contrario, el rastreador actualiza la posición del peatón más cercano de la lista a la posición actual, y guarda la posición anterior como registro de su trayectoria.
- (5) El rastreador elimina a los peatones de la lista si su posición no ha sido actualizada por el rastreador en β frames del proceso de rastreo.

Basándonos en nuestras observaciones, fijamos los valores de los parámetros $\alpha = 1$ m y $\beta = 5$ frames para todos nuestros estudios.

Predicción de posición y emisión sonora

El sistema predice las posiciones futuras de los peatones en las listas de seguimiento derivadas del proceso de seguimiento. El sistema utiliza las posiciones 3D de los peatones en el sistema de coordenadas de la cámara para predecir la velocidad relativa y la dirección de desplazamiento entre la maleta y cada peatón utilizando su posición actual y su trayectoria. Para mejorar la estabilidad de la medición de la posición de los peatones, el sistema compensa primero las rotaciones de la maleta (cámara) girando las posiciones detectadas de los peatones utilizando la información de la odometría 3D. A continuación, el sistema calcula la posición futura \hat{p}_{t+s} de cada peatón al cabo de s segundos utilizando los $N - 1$ puntos más recientes de su trayectoria de la siguiente manera:

procedimiento para

¹ \bar{o}_i
 actualizar una lista de seguimiento de los peatones
 detectados en cada fotograma:

- (1) El rastreador genera un conjunto de cuadros delimitadores a partir de la detección de peatones y calcula un conjunto de posiciones 3D

$$p_{\mu}(i, n) = \sum_{j=i-n+1}^n p_j \quad (1)$$

$$\hat{p}_{t+s} = \frac{p_{\mu}(t, N) - p_{\mu}(t - \frac{N}{2}, \frac{N}{2})}{\Delta t (p_t - \frac{N}{2}, p_t - \frac{3N}{2})} s + p_t \quad (2)$$

¹ <https://www.stereolabs.com/zed/>

Sea p_t la posición de un peatón en el sistema coordinado de cámaras en el tiempo t , $p_\mu(i, n)$ la media sobre n posiciones anteriores (de p_{i-n+1} a p_i), y $\Delta t(p_t, p_i)$ la diferencia en las marcas de tiempo entre p_t y p_i . El sistema calcula primero las dos posiciones medias (de $N - 1$ fotogramas antes a $N/2$ fotogramas antes, y de $N/2 - 1$ fotogramas antes al fotograma actual) (ecuación 1). A continuación, el sistema calcula la

entre las dos posiciones medias y predice la posición futura (ecuación 2). Con $N = 32$ se obtuvieron predicciones estables.

A continuación, el sistema predice el riesgo de una futura colisión basándose en todas las posiciones previstas para decidir si debe emitirse un sonido de alarma, como se indica en la figura 2. Se prevé una colisión cuando la trayectoria futura de un peatón cruza la "línea de emergencia" mostrada en la figura 2 (D). El sistema calcula la línea que une las posiciones actual y futura del peatón como predicción de la trayectoria futura prevista. A continuación, el sistema determina la intersección entre esta línea y la línea de emergencia. Si la intersección se encuentra dentro del alcance de la línea de emergencia, el riesgo de colisión se considera significativo y el sistema emite un sonido de aviso. Nótese que este cálculo no utiliza información sobre la altura de los peatones.

5 BBEEP: DISEÑO DE UNA POLÍTICA DE EMISIONES SÓLIDA MEDIANTE UN ESTUDIO DE OBSERVACIÓN

Estudiamos la respuesta de los peatones a las señales acústicas de advertencia con el fin de diseñar una política de emisión de sonido para nuestro sistema. Como se describe en Trabajos relacionados, se ha estudiado la percepción humana de los avisos acústicos de emergencia [12, 19, 32]. También se sabe cómo un estímulo visual puede hacer que los peatones redirijan su trayectoria [17]. Sin embargo, aún se sabe poco sobre cómo reacciona un peatón ante una señal acústica. Este conocimiento es importante para diseñar una política eficaz de emisiones sonoras para nuestro sistema de despeje de vías.

Realizamos un estudio observacional en un pasillo en el que se hizo que el sistema de maleta cerrada emitiera distintos tipos de alertas acústicas (pitidos). Se registraron las reacciones y trayectorias de los peatones, como se muestra en la figura 3. Se diseñó un conjunto de patrones de emisión de sonido que incluía varios tipos de sonido y una serie de tiempos. El sistema siguió a los peatones en el pasillo y emitió alertas utilizando estos patrones. Analizamos las trayectorias de los peatones para determinar qué patrones sonoros eran más eficaces para despejar el camino delante de la maleta. A partir de los resultados, diseñamos nuestra política de emisión de sonidos para evaluarla en un aeropuerto.

Se diseñaron 7 patrones sonoros (S1-S7), uno de los cuales (S7) era mudo para servir de referencia. Los 6 patrones no de referencia incluían combinaciones de 3 tipos de sonido diferentes y 2 tipos de tiempos de emisión.

Sonido Patrón	Tipo de sonido				Cronome traje <i>s</i>
	UL	BF	PD	IPI	
S1 S2	Alta	1000 Hz	0.1 s	0.1 s	5.0 s 2.5 s
S3 S4	Medio	400 Hz	0.1 s	0.1 s	5.0 s 2.5 s
S5 S6	Bajo	400 Hz	0.5 s	0.5 s	5.0 s 2.5 s
S7	Sin emisión sonora				N/A

Tabla 1: Patrones de emisión sonora. UL: nivel de urgencia, BF: frecuencia base, PD: duración del pulso, IPI: intervalo entre pulsos, y *s*: temporización (es decir, emite una alerta sonora teniendo en cuenta la posición prevista de los peatones transcurridos *s* segundos).

Tipos de sonidos de alerta. Utilizamos sonidos de alerta para representar 3 niveles distintos de urgencia percibida. La relación entre la urgencia percibida y los parámetros sonoros está bien documentada. Preparamos 3 tipos de pitidos con distintos niveles de urgencia: alto, intermedio y bajo. En concreto, variamos la frecuencia base, la frecuencia de pulso y el tono, tal y como se indica en la Tabla 1. Los valores que utilizamos se basan en estudios recientes sobre la percepción de la urgencia. Los valores que utilizamos se basan en investigaciones recientes sobre la urgencia del sonido [38, 41].

Temporización de la emisión sonora. También hemos utilizado distintos tiempos de emisión de sonidos para cada alerta sonora. El sistema cambia la temporización ajustando el parámetro de detección de colisiones *s* segundos. Si un sistema emitiera una alerta sonora inmediatamente antes de una colisión prevista con el usuario ciego (por ejemplo, $s < 1$), es posible que el peatón en cuestión no pueda evitar la colisión. Por otra parte, una alerta sonora emitida con demasiada antelación (por ejemplo, $s = \infty$) puede causar molestias e inconvenientes innecesarios y ser efectivamente improductiva.

Seleccionamos los valores de los parámetros $s_1 = 5,0$ y $s_2 = 2,5$ segundos. El valor de s_1 (5,0 s) representa el tiempo necesario para recorrer la distancia más lejana en el campo de detección (alrededor de 10 m) cuando el usuario ciego y un peatón se aproximan a una velocidad relativa de 4 km/h. También utilizamos s_2 (2,5 s) ajustado a la mitad de s_1 para definir un umbral más cercano.

Recogida y análisis de datos

En este estudio observacional se consideró el sistema de maleta cerrada con 7 patrones de emisión sonora colocado en un pasillo recto (figura 3). El sistema seguía a los peatones y predecía su

intersección con la línea de emergencia en tiempo real. El sistema también emitía los sonidos especificados por la política adoptada. Los siete patrones sonoros se utilizaron en ciclo.

Las observaciones se llevaron a cabo durante más de cuatro días, con 57 trayectorias para cada patrón (399 en total). Para cada patrón sonoro, el sistema registró una trayectoria e imágenes del pedestal más cercano con riesgo de colisión.



Figura 3: El sistema de maleta cerrada en un pasillo, equipado para emitir varios tipos y tiempos de pitidos.

Analizamos los conjuntos de datos registrados para determinar cómo afectaban las emisiones a las trayectorias. Ocasionalmente, algunas trayectorias faltaban o eran imprecisas debido a las limitaciones del procesamiento en tiempo real, como se describe en Seguimiento de peatones. Por tanto, realizamos un análisis posterior de las trayectorias utilizando las imágenes RGB y de profundidad grabadas para obtener trayectorias de peatones más precisas. Utilizamos OpenPose [8], un software de detección de cuerpos humanos basado en CNN, para detectar partes del cuerpo de los peatones a partir de las imágenes RGB. A continuación, determinamos la posición central de los cuerpos detectados en las imágenes de profundidad para obtener las posiciones 3D de los peatones. Realizamos el análisis para todas las imágenes grabadas. El ordenador portátil utilizado analizó las imágenes capturadas a una velocidad de 5 fps, es decir, una velocidad insuficiente para las notificaciones sonoras en tiempo real.

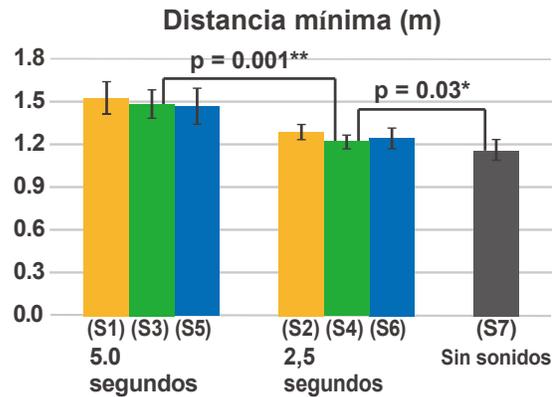
Medidas de evaluación

Medimos la "distancia mínima" entre la posición de la maleta y una trayectoria determinada para investigar la relevancia de distintos patrones de sonido a la hora de evitar colisiones. Una distancia mínima mayor puede interpretarse como un indicio de que el peatón ha evitado al usuario ciego por un margen cómodo. Estas distancias mínimas se determinaron a partir de las posiciones 3D devueltas por OpenPose.

Se consideraron tres hipótesis para los principales factores potenciales que influyen en las distancias mínimas: la presencia o ausencia de una emisión sonora (Hipótesis 1), el momento de la emisión sonora (Hipótesis 2) y el nivel de urgencia del sonido emitido (Hipótesis 3). Comprobamos estas hipótesis mediante una prueba de Kruskal-Wallis y una prueba U de Mann-Whitney con niveles de significación del 5 % para discernir las diferencias dentro de los patrones de sonido. También vimos los intervalos de confianza del 95 % para cada patrón.

Resultados

La figura 4 muestra la distancia mínima determinada para cada patrón. El patrón de línea de base muda (S7) arroja la



distancia mínima media más pequeña. La prueba de Kruskal-Wallis y la prueba U de Mann-Whitney, realizadas con un nivel de significación del 5%, revelaron que todos los patrones sonoros que no eran de línea de base (S1-S6) daban

Nivel de urgencia: ■ Alto ■ Medio ■ Bajo

Figura 4: Distancias mínimas. Las barras muestran los intervalos de confianza del 95 %. p : valor p de la prueba U de Mann-Whitney realizada sobre la distancia mínima (y * indican el 0,001 y 0,03 respectivamente).**

distancias mínimas más largas que S7. Este resultado valida la hipótesis 1. Basándonos en las pruebas estadísticas y en los intervalos de confianza del 95 %, también observamos que los patrones de emisión sonora con tiempos de 5 segundos (S1, S3 y S5) proporcionan distancias mínimas mayores que los patrones con tiempos de 2,5 segundos (S2, S4 y S6). Así pues, la hipótesis 2 también queda validada.

A continuación, comparamos los patrones de sonido con el mismo tiempo de emisión para evaluar la influencia de los niveles de urgencia. No se observaron diferencias estadísticamente significativas entre los patrones con tiempos de emisión de 5 segundos (S1, S3 y S5) o de 2,5 segundos (S2, S4 y S6). Por tanto, rechazamos la hipótesis 3.

A continuación resumimos nuestras conclusiones:

- Los avisos sonoros basados en la predicción de colisiones influyeron en los peatones que caminaban en dirección contraria a la maleta.
- La temporización de las emisiones sonoras también afectó a las trayectorias de los peatones. Los patrones con una temporización de 5 segundos desviaban trayectorias peatonales con mayor eficacia.
- El tipo de sonido de alarma no parece ser un factor significativo que afecte a las trayectorias de los peatones.

Diseño de la política de emisiones sonoras

Basándonos en las conclusiones anteriores, diseñamos una política de emisión de sonidos para BBeep (Figura 5), consistente en tres etapas de emisión de sonidos para evitar colisiones. El sistema emite los tres tipos de sonidos de alarma siguientes.

- (1) Un **pitido de baja urgencia** advierte del riesgo potencial de colisión entre el usuario ciego y los peatones en un plazo de 5 s. Este sonido se utilizó en nuestro estudio de observación como *sonido de baja urgencia* S5. Esperamos que esta señal permita a los peatones desviar su trayectoria lejos del usuario ciego para evitar la colisión.
- (2) Un **pitido de urgencia intermedia** indica un riesgo potencial de colisión en un plazo de 2,5 s. Este sonido se utilizó en nuestro

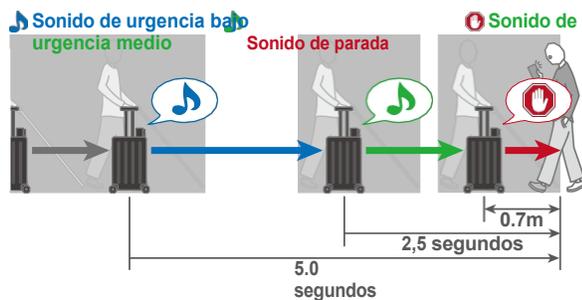


Figura 5: Política de emisión sonora.

estudio observacional como un *sonido de urgencia intermedia* S4. También esperamos que esta señal ayude a evitar la colisión.

- (3) Un **sonido de parada** indica un riesgo inminente de colisión con cualquier obstáculo (peatón, silla, pared, etc.) situado a menos de 70 cm. Esperamos que esta señal incite al usuario ciego a detenerse inmediatamente.

Elegimos los sonidos de urgencia intermedia y baja para nuestra política. Cuanto mayor era el nivel de urgencia de la señal, mayor era el grado de molestia de la alerta sonora. Sin embargo, observamos que el nivel de urgencia del sonido no afectaba a la trayectoria de los peatones que se aproximaban. Por lo tanto, seleccionamos dos alertas sonoras con niveles de urgencia y molestia más bajos. Utilizando dos tipos de pitidos, el usuario ciego puede saber si un peatón sigue acercándose o no. Además, para informar al usuario ciego de la presencia de un obstáculo, utilizamos un sonido de timbre completamente distinto de los pitidos. Este sonido de campana se emite siempre que el sistema detecta obstáculos situados a menos de 70 cm. En nuestra evaluación de usuarios, recomendamos que el usuario ciego deje de avanzar inmediatamente después de oír el sonido del timbre.

6 EVALUACIÓN DE USUARIOS EN EL MUNDO REAL

Nuestro principal objetivo era comprender la eficacia de BBeep para despejar el camino a los viajeros ciegos en espacios abarrotados. Para ello, realizamos un estudio real en el que 6 participantes ciegos (Tabla 2) navegaron por zonas abarrotadas en el Aeropuerto Internacional de Pittsburgh. En este estudio, comparamos BBeep con dos condiciones de referencia: una notifica únicamente al usuario ciego los riesgos de colisión, mientras que la otra no proporciona ninguna notificación.

Condiciones

Hemos dotado a nuestro sistema de asistencia para maletas de la capacidad de seguir a los peatones y predecir futuras colisiones. Basándonos en este sistema, preparamos tres interfaces diferentes:

Interfaz de altavoz (BBeep): Esta interfaz representa nuestro sistema propuesto, que emite tres tipos de sonidos (urgencia baja, urgencia media y sonidos de

ID	Género	Edad	Vista	Ayuda a la navegación
P1	F	70	A ciegas	Caña
P2	F	70	A ciegas	Caña
P3	M	65	A ciegas	Caña
P4	M	46	A ciegas	Perro (primario) y bastón
P5	M	42	A ciegas	Perro (primario) y bastón
P6	M	58	A ciegas	Caña

parada) para el usuario ciego y otros peatones a través de un altavoz que se monta en la parte superior de la maleta.

Tabla 2: Información demográfica de nuestros participantes

Interfaz de auriculares: Esta interfaz tiene el mismo comportamiento que BBeep, pero en lugar de utilizar un altavoz, proporciona sonidos sólo al usuario ciego utilizando auriculares conductores óseos (para evitar el bloqueo del sonido ambiental).

Interfaz sin sonido: El usuario también lleva la maleta, pero esta interfaz nunca emite sonido, lo que representa a un usuario ciego que navega por sí mismo sin ayudas adicionales.

Tareas

Seleccionamos varias puertas abarrotadas en las que los pasajeros esperaban el embarque en fila o en grupos. Se pidió a los participantes que caminaran en línea recta por el pasillo y atravesaran la multitud hasta llegar a un lugar determinado, donde el experimentador les pediría que se detuvieran (cada tarea tenía aproximadamente 20 metros). Los participantes sujetaban el asa de la maleta con una mano y utilizaban su bastón blanco con la otra (Figura 6). Su objetivo era atravesar la multitud con eficacia y evitar colisiones con otros peatones. Esta tarea nos permitió reproducir un escenario muy similar entre los distintos participantes y ensayos, lo que posibilitó una comparación justa entre las condiciones.

Procedimiento

Tras obtener el consentimiento informado de los participantes (aprobado por el IRB), los investigadores ofrecieron una visión general del estudio y describieron las tres interfaces. A continuación, se impartió a los participantes una breve sesión de formación (de 10 a 15 minutos) hasta que se familiarizaron con los sonidos de alarma del sistema y las interfaces. Aunque el volumen de la interfaz del altavoz era fijo para todos los participantes, podían ajustar el volumen de la interfaz de los auriculares para asegurarse de que era cómodo pero audible. Durante la sesión de entrenamiento, se explicó cómo sujetar la maleta, ya que afecta a la precisión de la predicción de colisiones. A continuación, se pidió a los participantes que recorrieran cinco rutas similares utilizando tres tipos de interfaces (la interfaz con altavoz y auriculares dos veces y la interfaz sin sonido una vez) en un orden contrapuesto. Se informó a los participantes de que un investigador caminaría detrás de ellos para garantizar su seguridad y la de los demás peatones (figura 6). También se les indicó que se detuvieran al escuchar el sonido de mayor urgencia (stop) para evitar colisionar con los peatones. El investigador



Figura 6: Estudio de usuarios en un aeropuerto internacional. Los participantes caminaron entre la multitud sujetando el asa de un sistema de asistencia con forma de maleta.

no intervenían a menos que: hubiera un riesgo inminente o una desviación del camino. Por ejemplo, en este último caso, el investigador les diría que ajustaran ligeramente su orientación. Además, en caso de que el camino estuviera bloqueado o el participante estuviera confundido, el investigador intervendría para ayudar al participante a pasar ese obstáculo inmediato.

Para observar la respuesta de los peatones al sistema, montamos una cámara GoPro en la parte superior de la maleta.

Métricas

Frecuencia de colisión inminente y frecuencia de riesgo de colisión. Para medir cuántos peatones tenían un riesgo inminente de colisión con el usuario ciego, definimos el número de peatones en un radio de 70 cm como la *Frecuencia de colisiones inminentes*. Además, medimos la *Frecuencia de riesgo de colisión*, que indica cuántos peatones corrían el riesgo de colisionar con el usuario ciego en un plazo de 5 s. En cada tarea, contamos las colisiones inminentes y el riesgo de colisión basándonos en los resultados de detección de peatones y en nuestros resultados de predicción de colisiones, respectivamente. Comparamos cuantitativamente las tres condiciones basándonos en un intervalo de confianza del 95% (Tabla 3). Además, comparamos las dos condiciones de sonido (interfaces de altavoz y auriculares) mediante una prueba de rangos con signo de Wilcoxon con niveles de significación del 1%.

Ratio de continuidad del riesgo. Esta métrica representa la proporción de peatones que presentaban riesgos potenciales de colisión y se mantuvieron en la trayectoria de los usuarios hasta alcanzar un riesgo inminente de colisión. Para calcular la métrica, dividimos la *Frecuencia de Colisión Inminente* por la *Frecuencia de Riesgo de Colisión*. Los valores más pequeños indican que el sistema reduce el riesgo de colisión entre el usuario ciego y los peatones. Realizamos el mismo análisis descrito para las métricas anteriores.

Entrevista posterior. Después de completar las tareas, pedimos a los participantes que valoraran una serie de frases utilizando ítems de Likert de 7 puntos (desde 1: totalmente en desacuerdo, 4: neutral, hasta 7: totalmente en desacuerdo).

de acuerdo). Las frases y un resumen de las respuestas se muestran en la Tabla 4. Por último, formulamos preguntas abiertas sobre las ventajas y dificultades de cada interfaz (altavoz y auriculares). También pedimos sugerencias para mejorar cada interfaz, y en qué escenarios utilizarían los participantes la interfaz de altavoz (BBeep).

Resultados

Evaluación cuantitativa. La tabla 3 muestra la frecuencia de colisiones inminentes, la frecuencia de riesgo de colisión y el coeficiente de continuidad del riesgo. No encontramos diferencias significativas entre las interfaces en cuanto a la frecuencia de riesgo de colisión ($p = 0,8$). Por otra parte, nuestro análisis reveló que la interfaz con altavoz dio lugar a un número significativamente menor ($p = 0,005$) de peatones con riesgo inminente de colisión con el usuario, en comparación con la interfaz con auriculares. Además, una diferencia significativa ($p = 0,009$) en la relación de continuidad del riesgo muestra que la interfaz con altavoz fue más eficaz para reducir el número de peatones que tenían un riesgo de colisión con el usuario.

La Tabla 4 muestra los resultados de seis preguntas de escala Likert. Cuatro participantes coincidieron en que la gente despejaba el camino cuando utilizaba la interfaz con altavoz en lugar de la interfaz con auriculares (Q1). En cambio, en las demás preguntas obtuvimos resultados similares entre las dos interfaces.

Observaciones de vídeo. Las grabaciones de vídeo nos permitieron analizar el comportamiento tanto del usuario ciego como de los peatones videntes, con el fin de complementar nuestras métricas cuantitativas. Observamos que los participantes despejaban el camino la mayoría de las veces tras percatarse de la presencia del usuario. Sin embargo, los participantes que utilizaban las interfaces con auriculares o sin sonido colisionaban a menudo (o corrían un riesgo inminente de colisión) con peatones que no se percataban de su presencia. En la mayoría de las ocasiones, los peatones estaban hablando en grupo o haciendo cola esperando para embarcar. En cambio, cuando se utilizaba la interfaz con altavoz, incluso en los escenarios mencionados los peatones oían el sonido y despejaban inmediatamente el camino al usuario ciego.

Sin embargo, hubo cinco excepciones en las que los peatones se acercaron en un radio de 70 cm del participante, lo que suponía un riesgo inminente de colisión. Los motivos fueron: (1) un peatón estaba usando auriculares y no oyó la notificación sonora; (2) un usuario ciego cambió repentinamente la dirección de su marcha; y (3) los peatones intentaron despejar el camino para el ciego, pero no tenían por dónde ir (por *ejemplo*, por estar contra la pared).

Comentarios cualitativos. En general, los participantes eran conscientes de que los demás peatones despejaban el camino cuando utilizaban BBeep, como ilustran sus comentarios: A1: "La ventaja del altavoz es [que] ellos [otros peatones] despejaban el camino" (P5); A2: "La gente se daba cuenta de que me acercaba y la gente se apartaba... dejándome el camino" (P4); y A3: "La mayor ventaja es que otras personas lo oían [la alerta sonora] y

Interfaz	Frecuencia del riesgo de colisión			Frecuencia de colisión inminente			Ratio de continuidad del riesgo		
	Media y DE	Inferior	Superior	Media y DE	Inferior	Superior	Media y DE	Inferior	Superior
Altavoz	6.67 ± 3.75	4.55	8.79	0.41 ± 0.76	0.00	0.85	0.08 ± 0.19	0.00	0.19
Auriculares	5.91 ± 2.25	4.64	7.19	2.00 ± 1.35	1.23	2.76	0.37 ± 0.25	0.22	0.51
Sin sonido	6.67 ± 2.05	5.03	8.30	3.00 ± 1.85	1.52	4.48	0.45 ± 0.21	0.28	0.63

Tabla 3: Evaluación cuantitativa de cada métrica. Presenta las medias y las desviaciones estándar (DE), así como los límites inferior y superior de los intervalos de confianza del 95 %.

No.	Pregunta	P1	P2	P3	P4	P5	P6	media	DE	Mediana
Q1	La gente despejaba el camino cuando yo utilizaba la interfaz del altavoz.	4	6	7	7	5	4	5.5	1.26	5.5
	La gente despejaba el camino cuando no utilizaba la interfaz del altavoz.	4	3	2	3	4	4	3.33	0.75	3.5
Q2	La interfaz del altavoz me ayudó a caminar cómodamente en los aeropuertos.	7	4	5	6	6	5	5.5	0.96	5.5
	La interfaz de los auriculares me ayudó a caminar cómodamente por los aeropuertos.	7	4	3	6	6	6	5.33	1.37	6
Q3	La interfaz de altavoz también es útil en lugares poco concurridos.	5	2	6	6	6	4	4.83	1.46	5.5
	La interfaz de auriculares también es útil en lugares menos concurridos.	5	2	4	5	6	4	4.33	1.25	4.5

Tabla 4: Ítems Likert (1: totalmente en desacuerdo a 7: totalmente de acuerdo) y resumen de las respuestas.

se movían para apartarse" (P3). Cuando utilizaban la interfaz de auriculares, los participantes consideraban que su principal ventaja era que era más silenciosa: A4: *"Es [la interfaz de auriculares] más privada"* (P1). Sin embargo, también tenían la percepción de que los peatones no despejaban el camino de la misma manera: A5: *"La gente no se da cuenta, así que tengo que decir algo para que despejen el camino... en comparación con el altavoz"* (P4); y A6: *"[La principal ventaja de la interfaz de auriculares es que] es silenciosa. [El*

el principal problema es que] no llamó la atención de nadie" (P2). Algunos participantes comentaron que la utilidad de la interfaz del altavoz podría depender del entorno. Los lugares en los que la utilizarían suelen ser espacios públicos abarrotados y zonas abiertas: A7: *"Es más útil en lugares más concurridos"* (P1); A8: *"Es útil en una tienda de comestibles, un centro comercial y otras zonas abiertas"* (P5). Sin embargo, también se señaló su utilidad en lugares menos concurridos: A9: *"La interfaz del altavoz también es útil en lugares menos concurridos, porque no pita cuando no hay gente. Así, puedo seguir aprovechándolo cuando hay alguien con riesgo de colisión"* (P3). Por el contrario, su uso en entornos más tranquilos se consideró inadecuado o que llamaba demasiado la atención: A10: *"En entornos tipo aeropuerto, probablemente utilizaría el altavoz, pero si estoy en una zona tranquila donde se espera que la gente esté callada,tal vez no usarlo"* (P4); y A11: *"No estoy de acuerdo con usar el altavoz inter- face en lugares supuestamente tranquilos como hospitales o bibliotecas, pero, en cualquier entorno público como aeropuertos, estaciones de tren o lo que sea, el altavoz siempre va a ser apropiado".* (P2).

Quando se les pidieron sugerencias, dos usuarios dijeron que BBeep no sólo debería emitir pitidos, sino también proporcionar información sobre el entorno: A12: *"Es más probable que utilice el altavoz, pero Todavía quiero oír lo que pasa a través de los auriculares"* (P3);

A13: "[En ambas interfaces] quiero entender lo que pasa. Hay gente delante de mí caminando, viniendo o de pie.

... [Recomiendo] una salida diferente. El altavoz avisará a las personas videntes. Los auriculares explicarán lo que está pasando". (P2).

7 DEBATE

Eficacia de BBEEP para evitar colisiones

El estudio con usuarios reales demostró que BBEEP era una herramienta eficaz para que los usuarios invidentes evitaran colisiones con peatones. Aunque el número de peatones con bajo riesgo de colisión (en los próximos 5 segundos) fue muy similar en las distintas condiciones, el número de peatones con riesgo inminente de colisión con el usuario fue significativamente inferior en el caso de BBEEP que en el de los auriculares. Estas dos condiciones proporcionan exactamente las mismas notificaciones sonoras, pero utilizan diferentes fuentes de salida (es *decir*, altavoz o auriculares). Este resultado indica que emitir sonido tanto al usuario ciego como a los peatones cercanos fue eficaz para despejar el camino al usuario, y que fue más eficaz que notificar sólo al usuario. La observación en vídeo de las tareas de navegación en el aeropuerto corrobora estos resultados. Los participantes atravesaron zonas abarrotadas de gente cerca de las puertas de embarque y se encontraron con frecuencia con peatones que no se percataban de su presencia. Cuando caminaban con BBEEP, los peatones videntes se percataban inmediatamente de la presencia del usuario, despejaban el camino y, en algunos casos, incluso pedían a sus compañeros que se movieran. Aunque no siempre eran conscientes del comportamiento de los usuarios videntes, los participantes tenían la percepción de que BBEEP era más eficaz que las alternativas, como demuestran sus valoraciones y comentarios (A1, A2 y A3).

Aunque la condición Auriculares no fue tan eficaz como BBEEP, los usuarios aprovecharon sus conocimientos sobre la colisión

riesgo con los peatones. Por ejemplo, P4 empezó a decir "¡Disculpe!" tras darse cuenta de que persistía el riesgo de colisión, mientras que otros participantes orientaron mejor el maletín para encontrar un camino sin riesgos.

Posibles escenarios para BBleep

Diseñamos cuidadosamente nuestra política de emisiones sonoras, teniendo en cuenta que la aceptación social era crucial para este enfoque. Además, realizar el experimento en el aeropuerto permitió a los participantes comprender cómo sería utilizar BBleep en el mundo real. Los comentarios de los participantes indican que es aceptable utilizar BBleep en espacios públicos abarrotados de gente, como puertos aéreos, estaciones de tren o centros comerciales (A7, A8, A10 y A11). De hecho, la sensación de comodidad de los usuarios al utilizar el sistema en forma de maleta (Q2) mostró resultados muy similares entre la interfaz Head-set y BBleep. En cambio, las opiniones de los participantes sobre el uso de ambas interfaces en lugares menos concurridos no son consensuadas (Q3). Aún así, algunos participantes ven ventajas en su uso, ya que no proporcionan notificaciones a menos que existan riesgos de colisión (A9). Aunque las zonas muy concurridas parecen apropiadas para utilizar BBleep, los participantes comentaron que no lo utilizarían en lugares muy tranquilos donde llamarían demasiado la atención o en lugares donde se supone que deben ser tranquilos, como hospitales o bibliotecas (A10 y A11).

Limitaciones y trabajo futuro

Reducción del número de emisiones sonoras. La principal ventaja de la interfaz de auriculares era su discreción, ya que no llama tanto la atención ni molesta a otras personas (A4 y A6). Sin embargo, el hecho de ser más privado repercutió significativamente en el rendimiento. Esta relación entre rendimiento y discreción sugiere que es importante investigar cómo reducir aún más el número de emisiones sonoras manteniendo su capacidad para despejar el camino a los usuarios ciegos. Por ejemplo, observamos que a veces se emitían notificaciones sonoras incluso cuando los peatones ya se habían percatado de la presencia del usuario ciego, pero no despejaban el camino inmediatamente. Futuras soluciones podrían considerar el uso de técnicas de seguimiento facial o estimación de la mirada [51] para evaluar si los peatones son conscientes del usuario ciego, reduciendo así el número de emisiones sonoras.

Aceptabilidad de BBleep por los peatones videntes. Para evaluar la aceptabilidad de BBleep, es importante investigar no sólo las impresiones de los usuarios ciegos, sino también las de los peatones videntes. Sin embargo, en este caso, reclutar de antemano a personas videntes nos impediría evaluar la capacidad de BBleep para ayudar a despejar el camino al usuario ciego. En el futuro, nos proponemos estudiar más a fondo las impresiones de las personas videntes con un diseño de estudio diferente.

Más allá del despeje del camino. Para evaluar el impacto de nuestro planteamiento, nos hemos centrado exclusivamente en la evitación de colisiones y en la capacidad de despejar el camino para el usuario ciego. En

Para ello, utilizamos rutas en línea recta y no incluimos dificultades de navegación adicionales que pudieran afectar a los resultados. Estas decisiones de diseño nos permitieron llevar a cabo un experimento más controlado, a pesar de realizarse en un escenario del mundo real. Sin embargo, atravesar de forma independiente entornos complejos como aeropuertos plantea retos adicionales importantes, como seguir una ruta concreta o conocer los puntos de interés (POI) de los alrededores.

Los participantes también mencionaron la necesidad de transmitir más información al usuario ciego, que quería conocer más detalles sobre su entorno (A12 y A13). Una posible ampliación es codificar la información de distancia (o urgencia) de forma continua en lugar de utilizar tres niveles predeterminados. Otra posibilidad es proporcionar al usuario información adicional útil para la orientación y la movilidad. En concreto, P2 y P3 sugirieron combinar el altavoz y los auriculares para que proporcionen información diferente al usuario. Sugirieron utilizar BBEEP tal cual, pero describir el entorno utilizando los auriculares óseo-conductores. En el futuro se podría investigar cómo combinar BBEEP con soluciones que proporcionen asistencia a la navegación giro a giro y/o transmitan información sobre PDI relevantes en las proximidades del usuario [1, 14, 42].

8 CONCLUSIÓN

Proponemos un sistema de maleta de asistencia, BBEEP, que pretende despejar el camino a los usuarios ciegos cuando caminan por espacios abarrotados, notificando tanto al usuario como a los peatones videntes los riesgos de colisión. Proporciona notificaciones sonoras sólo cuando es necesario, basándose en el seguimiento de los peatones y prediciendo su posición futura en tiempo real. Primero investigamos cómo transmitir eficazmente la información sonora a los peatones videntes y diseñamos la interfaz de notificación sonora de BBEEP. A continuación, realizamos un estudio de usuarios reales con discapacitados visuales en un aeropuerto. Los resultados mostraron que BBEEP reduce el número de situaciones de riesgo de colisión inminente en comparación con la notificación exclusiva al usuario invidente. Además, los usuarios consideraron aceptable y adecuado el uso de BBEEP en espacios públicos abarrotados, como aeropuertos, estaciones de tren o centros comerciales. Sin embargo, se mostraron más reticentes a utilizarlo en lugares en los que se supone que deben estar tranquilos. En el futuro, tenemos previsto ampliar nuestro método de predicción de colisiones utilizando un análisis de la atención basado en la visión para reducir el

número de emisiones sonoras innecesarias cuando los peatones ya se hayan percatado de la presencia del usuario ciego.

AGRADECIMIENTOS

Damos las gracias a la Autoridad Aeroportuaria del Condado de Allegheny y a todos los participantes en el estudio. Este trabajo ha sido patrocinado en parte por JST CREST (JPMJCR14E1), JST AIP-PRISM (JPMJCR18ZG), NSF NRI award (1637927), NIDILRR (90DPGE0003), Uptake (CMU ML for Social Good fund) y Shimizu Corporation.

REFERENCIAS

- [1] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hi-ronobu Takagi y Chieko Asakawa. 2016. NavCog: un asistente cognitivo de navegación para ciegos. En *Proc. ACM Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, 90-99. <https://doi.org/10.1145/2935334.2935361>
- [2] Tomohiro Amemiya, Jun Yamashita, Koichi Hirota y Michitaka Hirose. 2004. Guías virtuales para sordociegos: A real-time way-finder by verbal-nonverbal hybrid interface and high-density RFID tag space. En *Proc. IEEE Conference on Virtual Reality (VR '04)*. IEEE, 165-287. <https://doi.org/10.1109/VR.2004.1310070>
- [3] Mauro Ávila Soto, Markus Funk, Matthias Hoppe, Robin Boldt, Katrin Wolf y Niels Henze. 2017. DroneNavigator: Uso de cuadricópteros con correa y de flotación libre para la navegación de viajeros con discapacidad visual. En *Proc. ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*. ACM, 300-304. <https://doi.org/10.1145/3132525.3132556>
- [4] JM Bengamin. 1973. El nuevo bastón láser C-5 para ciegos. En *la Conferencia sobre la prótesis electrónica, 1973*.
- [5] BB Blasch, SJ LaGrow y WR De l'Aune. 1996. Three aspects of coverage provided by the long cane: Object, surface, and foot-placement *r e v i e w*. *Journal of Visual Impairment and Blindness* 90 (1996), 295-301.
- [6] GP Boucek Jr, James E Veitengruber y Wayne D Smith. 1977. *Aircraft alerting systems criteria study. Volume II. Human factors guidelines for aircraft alerting systems*. Informe técnico. BOEING COMMERCIAL AIRPLANE CO SEATTLE WA.
- [7] Michael Brock y Per Ola Kristensson. 2013. Supporting blind navigation using depth sensing and sonification. En *Proc. ACM International Joint Conference on Pervasive and ubiquitous computing adjunct publication (UbiComp '13)*. ACM, 255-258. <https://doi.org/10.1145/2494091.2494173>
- [8] Zhe Cao, Tomas Simon, Shih-En Wei y Yaser Sheikh. 2017. Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. En *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '17)*. IEEE, 1302-1310. <https://doi.org/10.1109/CVPR.2017.143>
- [9] Ping-Ling Chen, Wafaa Saleh y Chih-Wei Pai. 2017. Texting and walking: a controlled field study of crossing behaviours and inattentive blindness in Taiwan. *Behaviour & Information Technology* 36, 4 (2017), 435-445. <https://doi.org/10.1080/0144929X.2016.1240234>
- [10] Sakmongkon Chumkamon, Peranitti Tuvaphanthaphiphat y Phongsak Keeratiwintakorn. 2008. A blind navigation system using RFID for indoor environments. En *Proc. IEEE International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON '08)*, Vol. 2. IEEE, 765-768. <https://doi.org/10.1109/ECTICON.2008.4600543>
- [11] Karen Duarte, José Cecilio y Pedro Furtado. 2014. Guiar fácilmente a los ciegos: Proporcionando información y navegación-smartnav. En *Wireless Internet*. Springer, 129-134. https://doi.org/10.1007/978-3-319-18802-7_18
- [12] Judy Edworthy, Sarah Loxley e Ian Dennis. 1991. Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human factors* 33, 2 (1991), 205-231. <https://doi.org/10.1177/001872089103300206>
- [13] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris y Eelke Folmer. 2012. The user as a sensor: navigating users with visual impairments in indoor spaces using tactile landmarks. En *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, 425-432. <https://doi.org/10.1145/2207676.2207735>
- [14] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris y Eelke Folmer. 2013. Sistemas de navegación humana en interiores: A survey. *Interacting with Computers* 25, 1 (2013), 21-33. <https://doi.org/10.1093/iwc/iws010>
- [15] José Faria, Sérgio Lopes, Hugo Fernandes, Paulo Martins y João Barroso. 2010. Bastón blanco electrónico para la asistencia a la navegación de personas ciegas. En *Proc. IEEE Conferencia sobre el Congreso Mundial de Automatización (WAC '10)*. IEEE, 1-7.
- [16] Vitor Filipe, Filipe Fernandes, Hugo Fernandes, António Sousa, Hugo Paredes y João Barroso. 2012. Sistema de apoyo a la navegación a ciegas basado en Microsoft Kinect. *Procedia Computer Science* 14 (2012), 94-101. <https://doi.org/10.1016/j.procs.2012.10.011>
- [17] Masahiro Furukawa, Hiromi Yoshikawa, Taku Hachisu, Shogo Fukushima y Hiroyuki Kajimoto. 2011. "Campo de vección" para el control del tráfico peatonal. En *Proc. ACM International Conference on Augmented Human (AH '11)*. ACM, 19. <https://doi.org/10.1145/1959826.1959845>
- [18] Thomas Gallagher, Elyse Wise, Binghao Li, Andrew G Dempster, Chris Rizos y Euan Ramsey-Stewart. 2012. Indoor positioning system based on sensor fusion for the blind and visually impaired. En *Proc. IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN '12)*. IEEE, 1-9. <https://doi.org/10.1109/IPIN.2012.6418882>
- [19] Christian González, Bridget A Lewis, Daniel M Roberts, Stephanie M Pratt y Carryl L Baldwin. 2012. Percepción de urgencia y molestia de las alertas auditivas en un contexto de conducción. 56, 1 (2012), 1684-1687. <https://doi.org/10.1177/1071181312561337>
- [20] Ellen C Haas y Judy Edworthy. 1996. Designing urgency into auditory warnings using pitch, speed and loudness. *Computing & Control Engineering Journal* 7, 4 (1996), 193-198. <https://doi.org/10.1049/cc:19960407>
- [21] Juan David Hincapié-Ramos y Pourang Irani. 2013. CrashAlert: mejora de la alerta periférica para la interacción móvil ojos ocupados mientras camina. En *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, 3385-3388. <https://doi.org/10.1145/2470654.2466463>
- [22] Hsieh-Chang Huang, Ching-Tang Hsieh y Cheng-Hsiang Yeh. 2015. Un sistema de detección de obstáculos en interiores utilizando información de profundidad y el crecimiento de la región. *Sensors* 15, 10 (2015), 27116-27141. <https://doi.org/10.3390/s151027116>
- [23] Andreas Hub, Joachim Diepstraten y Thomas Ertl. 2004. Diseño y desarrollo de un sistema de navegación en interiores e identificación de objetos para ciegos. En *Proc. ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '04)*. ACM, 147-152. <https://doi.org/10.1145/1029014.1028657>
- [24] Kiyohide Ito, Makoto Okamoto, Junichi Akita, Tetsuo Ono, Ikuro Gyobu, Tomohito Takagi, Takahiro Hoshi y Yu Mishima. 2005. CyARM: un dispositivo de ayuda alternativo para personas ciegas. En *Proc. ACM CHI Extended Abstracts on Human Factors in Computing Systems (CHI '05)*. ACM, 1483-1488. <https://doi.org/10.1145/1056808.1056947>
- [25] Beomjoon Kim y Joelle Pineau. 2016. Socially Adaptive Path Planning in Human Environments Using Inverse Reinforcement Learning. *International Journal of Social Robotics* 8, 1 (2016), 51-66. <https://doi.org/10.1007/s12369-015-0310-2>
- [26] Jee-Eun Kim, Masahiro Bessho, Shinsuke Kobayashi, Noboru Koshizuka y Ken Sakamura. 2016. Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. En *Proc. ACM Annual Symposium on Applied Computing (AC '16)*. ACM, 604-611. <https://doi.org/10.1145/2851613.2851716>
- [27] Kris M. Kitani, Brian D. Ziebart, James Andrew Bagnell y Martial Hebert. 2012. Activity Forecasting. En *Proc. IEEE European Conference on Computer Vision (ECCV '12)*, Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, y Cordelia Schmid (Eds.). Springer Berlin Heidelberg, 201-214. <https://doi.org/10.1016/B978-0-12-809276-7.00014-X>
- [28] Vladimir Kulyukin, Chaitanya Gharpure, John Nicholson y Grayson

Osborne. 2006. Orientación asistida por robots para discapacitados visuales.

- en entornos interiores estructurados. *Robots autónomos* 21, 1 (2006), 29–41. <https://doi.org/10.1007/s10514-006-7223-8>
- [29] Bing Li, J Pablo Muñoz, Xuejian Rong, Jizhong Xiao, Yingli Tian y Ariés Arditi. 2016. ISANA: wearable context-aware indoor assistive En navegación con evitación de obstáculos para ciegos. En *Proc. IEEE EuropeanProc Automatización. IEEE, Conference on Computer Vision (ECCV '16)*. Springer, 448-462. https://doi.org/10.1007/978-3-319-48881-3_31
- [30] Shachar Maidenbaum, Shlomi Hanassy, Sami Abboud, Galit Buchs, Daniel-Robert Chebat, Shelly Levy-Tzedek, and Amir Amedi. 2014. El "EyeCane", una nueva ayuda electrónica de viaje para ciegos: Tecnología comportamiento y aprendizaje rápido. *Neurología reparadora y neurociencia* 32, Enhance 6 (2014), 813-824.
- [31] Roberto Manduchi, Sri Kurniawan y Homayoun Bagherinia. 2010. Blind guidance using mobile computer vision: Un estudio de usabilidad. En *Proc. applying ACM SIGACCESS Conference on Computers and accessibility (IEEE Trans- '10)*. ACM, 241-242. <https://doi.org/10.1145/1878803.1878851>
- [32] Dawn C Marshall, John D Lee y P Albert Austria. 2007. Alertas para sistemas de información a bordo: Annoyance, urgency, and appropriateness. *Human factors* 49, 1 (2007), 145-157. <https://doi.org/10.1518/American para 001872007779598145>
- [33] Edward Marshall y Sue Baker. 1995. Alarms in nuclear power plant control rooms: current approaches and future design. En *Human factors2018 Proc. IEEE en diseño de alarmas*. Taylor & Francis, Inc., 183-191. *reconocimiento de patrones*.
- [34] Christina Meredith y Judy Edworthy. 1995. Are there too many alarms in the unidad de cuidados intensivos? Una visión general de los problemas. *Jour- Wu*. 2016. Co-robotic nal of advanced nursing 21, 1 (1995), 15-20. navegación para discapacitados visuales. *IEEE* 001872089303500408
- [35] Madoka Nakajima y Shinichiro Haruyama. 2012. Navegación en interiores. sistema para personas con discapacidad visual que utiliza comunicación por luz geomagnética compensada y basada en cámara. En *Proc. IEEE Internationalmobile mapping of un- Conferencia sobre Comunicaciones en China (ICCC '12)*., 524-529. *International Conference on* <https://doi.org/10.1109/ICCCChina.2012.6356940>
- [36] Helen Petrie, Valerie Johnson, Thomas Strothotte, Andreas Raab, Steffi Fritz, y Rainer Michel. 1996. MoBIC: Designing a travel aid for personas ciegas y de edad avanzada. *The Journal of Navigation*49, 1 (1996), 45-52. *gaze estima- Pattern Analysis and Machine Intelligence (Transacciones del IEEE sobre análisis de patrones e inteligencia artificial)*.
- [37] Huy-Hieu Pham, Thi-Lan Le y Nicolas Vuillermé. 2016. Real-time obstacle detection system in indoor environment for the visually impaired utilizando el sensor kinect de microsoft. *Journal of Sensors* 2016 (2016). <https://doi.org/10.1155/2016/3754918>
- [38] Ioannis Politis, Stephen A Brewster y Frank Pollick. 2014. Evaluating multimodal driver displays under varying situational urgency. En *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, 4067-4076. <https://doi.org/10.1145/2556288.2556988>
- [39] Lisa Ran, Sumi Helal y Steve Moore. 2004. Drishti: un sistema y servicio integrado de navegación para invidentes en interiores y exteriores. En *Proc. the Second IEEE Annual Conference on Pervasive Computing and Communications (PerCom '04)*. IEEE, 23-30. <https://doi.org/10.1109/PERCOM.2004.1276842>
- [40] Joseph Redmon y Ali Farhadi. 2017. YOLO9000: mejor, más rápido, más fuerte. En *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '17)*. IEEE, 6517-6525. <https://doi.org/10.1109/CVPR.2017.690>
- [41] Shadan Sadeghian Borojeni, Susanne CJ Boll, Wilko Heuten, Heinrich H Bühlhoff y Lewis Chuang. 2018. Feel the Movement: Real Motion Influences Responses to Take-over Requests in Highly Automated Vehicles. En *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 246. <https://doi.org/10.1145/3173574.3173820>
- [42] Daisuke Sato, Uran Oh, Kakuya Naito, Hironobu Takagi, Kris Kitani y Chieko Asakawa. 2017. Navcog3: Una evaluación de un asistente de navegación en interiores ciego basado en teléfonos inteligentes con características semánticas en un entorno a gran escala. En *Proc. ACM Ordenadores y accesibilidad (ASSETS '17)*. ACM, 270-279. <https://doi.org/10.1145/3132525.3132535>
- [43] Shraga Shoval, Johann Borenstein y Yoram Koren. 1994. Mobile robot obstacle avoidance in a computerized travel aid for the blind. *IEEE Conferencia Internacional sobre Robótica y* 2023-2028.
- [44] Shraga Shoval, Johann Borenstein y Yoram Koren. 1998. The Navbelt-A computerized travel aid for the blind based on mobile robotics technology. *IEEE Transactions on Biomedical Engineering* 45, 11 (1998), 1376-1386. <https://doi.org/10.1109/10.725334>
- [45] Yu-Chih Tung y Kang G Shin. 2018. Use of Phone Sensors to Seguridad de los peatones distraídos. *IEEE Transactions on Mobile Computing*, 17, 6 (2018), 1469-1482. <https://doi.org/10.1109/TMC.2017.2764909>
- [46] Iwan Ulrich y Johann Borenstein. 2001. The GuideCane-ASSETSmobile robot technologies to assist the visually impaired. *actions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 31, 2 (2001), 131-136.
- [47] William R Wiener, Richard L Welsh, and Bruce B Blasch. 2010. *daciones de orientación y movilidad*. Vol. 1. Fundación the Blind.
- [48] Takuma Yagi, Karttikeya Mangalam, Ryo Yonetani y Yoichi Sato. Futura localización de personas en videos en primera persona. En *Conferencia internacional sobre visión por ordenador y (CVPR '18)*.
- [49] Cang Ye, Soonhac Hong, Xiangfei Qian y Wei <https://doi.org/10.1177/cane:> Una nueva ayuda robótica a la *Systems, Man, and Cybernetics Magazine* 2, 2 (2016), 33-42. <https://doi.org/10.1109/MSMC.2015.2501167>
- [50] Limin Zeng, Markus Simros y Gerhard Weber. 2017. Detección electronic travel aids support for cognitive knownspaces. En *Proceedings of the 19th Interacción Persona-Ordenador con Dispositivos y Servicios Móviles HCI '17)*. ACM, 8. <https://doi.org/10.1145/3098279.3098563>
- [51] Xucong Zhang, Yusuke Sugano, Mario Fritz y Andreas Bulling. 2017. Real-world dataset and deep appearance-based <https://doi.org/10.1145/191028.191051tion>. *IEEE Transactions on* (2017). <https://doi.org/10.1109/TPAMI.2017.2778103>

SIGACCESS Conference on

[Ver estadísticas de publicación](#)