

SISTEMA DE NAVEGACIÓN Y EXPLORACIÓN VISUAL ARTICULADO POR TÉCNICAS DE INTELIGENCIA ARTIFICIAL, PARA LA MOVILIDAD DE PERSONAS EN CONDICIÓN DE DISCAPACIDAD VISUAL

Investigador Principal:
Juan Manuel Aldana Porras
Ingeniero de Sistemas
Correo: juan.aldana@unad.edu.co

Líder del Semillero CODESIST
Nilson Albeiro Ferreira Manzanares
Ingeniero de Sistemas
Especialista en Pedagogía para el Desarrollo
del Aprendizaje Autónomo
Master of Art in Education- Online Education
Correo: nilson.ferreira@unad.edu.co

Co-investigador:
John Fredy Montes Mora
Ingeniero de Sistemas con énfasis en
Telecomunicaciones
Especialista en Informática y Telemática
Candidato a Magister en E-Learning UNAB-
UOC
Correo: john.montes@unad.edu.co

Grupo de Investigación GIDESTEC - Semillero CODESIST Universidad Nacional Abierta y a Distancia Ibagué - Tolima

RESUMEN

La vida de una persona en condición de discapacidad visual se ve afectada por muchos factores, uno de ellos es la independencia, la dificultad para movilizarse de manera segura por la ciudad es uno de los principales obstáculos en su inclusión laboral y social, y aunque sus necesidades específicas suelen variar dependiendo del tipo de discapacidad visual, su edad y su habilidad adquirida para llevar a cabo tareas rutinarias, la ciudad de Ibagué no está hecha para favorecer a este tipo de población, que tiene que lidiar con los mismos problemas de cualquier ciudadano pero con la desventaja de no poder ver, teniendo que enfrentar todo tipo de obstáculos como huecos, postes mal ubicados, espacio público limitado, desniveles, escaleras y salientes en los andenes, propios de la falta de control de los entes gubernamentales y la falta de conciencia de las propios ciudadanos, estos obstáculos no pueden ser sorteados simplemente con el uso del bastón, esto termina obligándolos a ir siempre guiados por un acompañante, lo cual limita de manera significativa su independencia. Por ello se propone la implementación un dispositivo mediado por técnicas de inteligencia artificial, en el cual se agrupan gran cantidad de conceptos que van desde la visión computacional hasta las máquinas de vectores

de soporte, este algoritmo se ejecutara en un dispositivo que cuenta con un par de cámaras digitales, que harán las veces de sistema estereoscópico y serán las encargadas de tomar la información del medio, mapear tridimensionalmente el entorno, procesar la información, procesar la información a manera de lenguaje natural y transmitido al usuario en forma de comandos de voz que le permitirán sortear los obstáculos más comunes presentes en la ciudad.

PALABRAS CLAVE

Inteligencia, Artificial, Discapacidad, Visual, Navegación.

PROBLEMA

La movilidad dentro de las ciudades en Colombia de por sí ya es un problema propio de la poca visión de nuestros dirigentes y de la falta de una infraestructura adecuada, que con el tiempo, termina colapsando y generando problemas más grandes de lo que debería, si nos contextualizamos un poco, este tipo de problemas de movilidad se multiplican de manera exponencial para las personas con algún tipo de discapacidad, en especial para las personas con problemas visuales; y a pesar de que las personas con este tipo de discapacidad adquieren durante su vida una serie de habilidades que les permiten

desplazarse siguiendo ciertos lineamientos que hasta determinado punto están pensadas para que lo hagan de manera más segura, una de las ayudas técnicas más usada es el bastón, que con el tiempo y la práctica permite realizar tareas rutinarias en entornos conocidos, inclusive estableciendo rutas a lugares que se frecuentan comúnmente (el trabajo, la casa de un amigo, el supermercado, etc...) con cierta seguridad, el problema radica en que este entorno (por fuera de su casa) es cambiante y aparecen nuevos obstáculos en cualquier momento, como nuevas construcciones, vehículos mal ubicados, motos estacionadas en el andén, postes y templete (que a pesar de la práctica y de conocer el recorrido siempre son un obstáculo difícil de sortear), ventanas abiertas, o cualquier tipo de salientes que se encuentren por encima de radio de acción del bastón, huecos, y cualquier tipo de vehículo en movimiento, en especial aquellos que generen poco ruido, como los eléctricos y que son difíciles de detectar de forma auditiva, lo que hace que a larga estas habilidades se queden cortas frente a las necesidades reales de cada uno de ellos.

FORMULACIÓN DEL PROBLEMA

¿Es posible crear un dispositivo mediado por técnicas de inteligencia artificial que sirva como

medio de navegación urbana para personas con problemas de discapacidad visual, que permita detectar obstáculos tridimensionales en entornos complejos y cambiantes como las calles de la ciudad de Ibagué?

OBJETIVOS DE LA INVESTIGACIÓN GENERAL

✓ Implementar un dispositivo capaz de reconocer obstáculos tridimensionales en ambientes cambiantes como las calles de la ciudad de Ibagué, con un gran nivel de precisión y resistencia a cambios de ambiente y entornos.

OBJETIVOS ESPECÍFICOS

✓ Crear un detector de objetos tridimensionales a partir de un mapa de disparidad capturado por el sistema estereoscópico.

✓ Delimitar objetos y extraer información relevante a partir del mapa de disparidad generado, obteniendo solo elementos de interés particular (obstáculos).

✓ Encontrar las condiciones ideales para que el dispositivo pueda desenvolverse en las calles de la ciudad teniendo un alto grado de precisión y robustez a cambios de entornos.

✓ Lograr que el dispositivo tenga un porcentaje de precisión por encima del 90% en la detección de obstáculos genéricos.

ANTECEDENTES

Durante mucho tiempo se han venido desarrollando dispositivos de ayuda para personas en condición de discapacidad visual, la mayoría de ellos están relacionados con tecnologías que utilizan sensores ultra sónicos, infrarrojos o laser, pero ninguno de ellos ha sido capaz de tener éxito comercial, esto debido en parte a tres grandes aspectos, una interfaz de usuario que es poco clara y en muchos casos requiere entrenamiento por parte del usuario, un bajo porcentaje de detección sumado a unos costos elevados y una estética poco cuidada. Por otro lado el campo de la visión artificial en los últimos años se ha visto beneficiado por los

constantes avances de la tecnología y poder de procesamiento, acompañado de numerosas iniciativas en diversas universidades a nivel global que han impulsado el desarrollo de esta área, algunos de estos esfuerzos se han orientado a desarrollar ayudas técnicas para personas en condición de discapacidad visual, en este apartado revisaremos algunas de ellas. La mayoría de los estudios encontrados en esta área están directamente relacionados con la visión estereoscópica de una u otra manera, así pues se pueden destacar los estudios de la Universidad de Wollongong en Australia con su proyecto ENVIS, Electro Neural Vision System, el TVS (Tactile Vision System) desarrollado por la Universidad de Arizona y el proyecto ATAD, realizado por la universidad Carlos III de Madrid, Figura 1.



Figura 1. De Izquierda a derecha, ENVIS - Universidad de Wollongong, TVS (Tactile Vision System) Universidad de Arizona, ATAD, Universidad Carlos III de Madrid

MARCO TEÓRICO

Visión Computacional

La visión computacional es una rama de la inteligencia artificial dedicada a capturar información del mundo real y darle un sentido propio que las máquinas puedan entender, esto con el fin de dotarlas de información relevante de su entorno brindándoles la capacidad de tomar decisiones en base a ello, de una forma más estricta podríamos decir que la visión artificial o comprensión de imágenes

describe la deducción automática de la estructura y propiedades de un mundo tridimensional, posiblemente dinámico, bien a partir de una o varias imágenes bidimensionales de ese mundo (Nalwa, 1993)

Librerías de Visión Computacional OpenCV

OpenCV es un conjunto de librerías de código abierto dedicadas a la visión computacional, con una gran cantidad de algoritmos enfocados a la solución de problemas relacionados con esta área, opencv está estructurado de manera modular, lo que significa que incluye gran cantidad de librerías compartidas y estáticas, que ofrecen operaciones básicas de procesamiento de imágenes, análisis estructural, análisis de movimiento, reconocimiento del modelo, reconstrucción 3d entre otras lo cual es de vital importancia para el desarrollo del programa, ya que son las bases de trabajo e implementación del algoritmo final.

Sistema Estereoscópico

La estimación de profundidad de un objeto presente en una escena a partir de un sistema estéreo es el punto de partida del algoritmo desarrollado, para ello utilizamos dos (o más) imágenes separadas en el espacio, tal y como las proporcionan los ojos de los humanos situados en la parte delantera de la cara. Puesto que un elemento dado de la escena estará en diferente lugar en relación con el eje z de cada plano de imagen, si superponemos las dos imágenes habrá una disparidad en la localización del elemento en las dos imágenes. (Norvig & Russell, 2004), podemos definir la disparidad como la diferencia de desplazamiento en el eje x de un punto específico presente en las dos imágenes, que se encuentra perfectamente alineado en el eje y

Máquina de Vectores de Soporte

Las máquinas de vectores de soporte o SVM (support vector machine, por su siglas en inglés) es un método de clasificación basado en la minimización del riesgo estructural (SRM) de la teoría del aprendizaje estadístico, que funcionan como clasificadores en multitud de ámbitos, para el contexto general de este

trabajo se complementaran con los vectores de características obtenidos a partir de del modelo HOG, mapeando los puntos de entrada a un espacio de características de una dimensión mayor, para luego encontrar el hiper plano que los separe y maximice el margen entre las clases, figura 2.

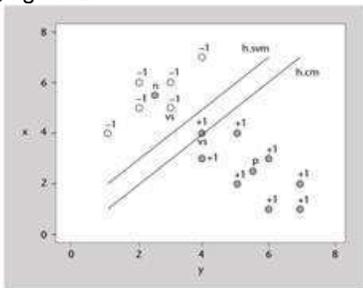


Figura 1. Ejemplo Maquina de Vectores de Soporte (Colmenares, 2009)

Descriptor de Imagen

Los descriptores de imagen son un método de representación matemática de contenidos específicos dentro de una imagen, en un concepto más simple puede verse como la forma de cuantificar matemáticamente un elemento presente en una imagen.

METODOLOGÍA

Para el desarrollo de este trabajo y teniendo en cuenta que es una investigación de tipo aplicada, se hace necesario construir un dispositivo de pruebas que sirva como puente inicial para llevar a cabo todo el trabajo de campo, (toma de muestras, mediciones de rendimiento, etc..) para lo cual y luego de realizar una búsqueda en el mercado teniendo siempre en cuenta que uno de los objetivos de esta tesis es hacer que el dispositivo final sea económico y pueda ser fácilmente adquirido, se determinaron los componentes de este buscando el equilibrio entre costo y rendimiento.

Dispositivo de Pruebas

Como plataforma de trabajo y pruebas se seleccionó cuidadosamente el hardware y software para llevar a cabo todo el proceso

investigativo, por tal motivo las herramientas libres surgieron como mejor alternativa, no solo por su estabilidad, robustez y gran cantidad de literatura disponible, sino también por ser comunes en el ámbito de la inteligencia artificial, igualmente cabe resaltar que el proyecto en su totalidad se realizó con software libre:

Lenguaje de Programación: C++

Librerías: OpenCV 2.4.12

Entorno de desarrollo: GNU/Linux

IDE: Code Blocks

El hardware está conformado por una Raspberry pi 2, dos cámaras digitales Genius montadas en un marco acrílico y una batería externa de 5000 mah a 5V.

RESULTADOS

Para evaluar el rendimiento general de detector de obstáculos se diseñó una prueba partiendo de escenarios reales y con cierta complejidad, la idea es evaluar el dispositivo de dos formas diferentes, la primera como detector de obstáculos tridimensionales, y la segunda con un valor agregado utilizando la máquina de vectores de soporte para convertirlo en un detector de objetos tridimensionales capaz de identificar específicamente personas además de objetos genéricos; para realizar la primera evaluación se estableció un recorrido de 2 kilómetros a lo largo del centro de la ciudad de Ibagué.

Total Reales Positivos = 328

Total Reales Negativos = 31

Total Falsos Positivos = 51 (detecciones erróneas a causa del ruido)

Tasa de Detección Global = Reales Positivos/Objetos GroundTruth

Tasa de Detección Global = $328/359 = 0,91$

Tasa de Error = $1 - \text{Tasa de Detección Global} = 0,09$

Detecciones Totales = Total reales positivos + Total reales negativos + Falsos Positivos = $328+31+51 = 410$

Exactitud = $\text{Reales Positivos} / \text{Detecciones Totales} = 328/410 = 0,80$

Presión = $\text{Reales Positivos} / \text{Reales Positivos} + \text{Falsos Positivos} = 328/328+51 = 0,86$

De estos datos podemos resaltar la tasa de detección que se encuentra por encima del 90%, la exactitud al momento de encontrar los obstáculos que esta sobre el 80% y la precisión que esta sobre el 86%.

REFERENCIAS

Alba, F, Castejón, M, González , A, Martínez de Pisón, J, Ordieres, J, Pernía, A, y Vergar, E (2006), Técnicas y algoritmos básicos de visión artificial, Universidad de La Rioja, Servicio de Publicaciones.

Alvarado, P, (2013), Vision por Computador, Costa Rica, Tecnológico de Costa Rica, Recuperado de:

<http://www.ie.itcr.ac.cr/siplab/index.php/PabloAlvarado/MScVisi%C3%B3nPorComputador>

Csurka, G., Dance, C., Fan, L., Willamowski, J., & Bray, C. (2004, May). Visual categorization with bags of keypoints. In Workshop on statistical learning in computer vision, ECCV (Vol. 1, No. 1-22, pp. 1-2).

Johnson L. y Higgins C. (2006) "A navigation aid for the blind using tactile-visual sensory substitution." En International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6289–6292.

Khan A., Moideen F., Lopez J., Khoo W., y Zhu Z., "KinDetect: Kinect Detecting Objects".

En Computers Helping People with Special Needs, Vol. LNCS7383, pp. 588–595

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In Computer vision, 1999. The proceedings of the seventh IEEE international conference on (Vol. 2, pp. 1150-1157). Ieee.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), 91-110.

Meers S. y Ward K. (2005) "A Substitute Vision System for Providing 3D Perception and GPS Navigation via Electro-Tactile Stimulation". En International Conference on Sensing Technology, pp. 551–556.

Nalwa, V.S (1993), A Guided tour to computer vision. Michigan, Estados Unidos, Addison-Wesley.

Organización Nacional de Ciegos Españoles.(2011), Discapacidad Visual y Autonomía personal, Madrid España, IRC

Peris, M (2014), Opencv: Stereo Camera Calibration, Recuperado de: <http://blog.martinperis.com/2011/01/opencv-stereo-camera-calibration.html>

Raducanu, B., Salas, J., Terven, J.,(2013), Estado del Arte en Sistemas de Visión Artificial para Personas Invidentes, Sociedad Mexicana de Inteligencia Artificial, Año V, Vol I: 20-30

Real, P, Jimenez, M, (2013), Curso Procesamiento De Imágenes Digitales, Universidad de Sevilla, Recuperado de: <http://grupo.us.es/gtocom/pid/pid10/doc.htm>

Revuelta Sanz, P. et al. (2013): "ATAD: Una Ayuda Técnica para la Autonomía en el Desplazamiento", Revista Española de Discapacidad, I (2): 143-154.

Russell, S, Norvig, P.(2004), Inteligencia Artificial. Un Enfoque Moderno. Segunda Edición, Madrid, España, Pearson Educación S.A.

Hernández García, R., García Reyes, E., Ramos Cózar, J., & Guil Mata, N. (2014). Modelos de representación de características para la clasificación de acciones humanas en video: estado del arte. Revista Cubana de Ciencias Informáticas, 8(4), 21-51.

Valveny, E., Varnell, M., Lopez, A.,(2015), Detección de Objetos, Universidad Autónoma de Barcelona, Recuperado de:<https://www.coursera.org/learn/deteccion-objetos/home/welcome>

Velazquez R., Fontaine E., y Pissaloux E. (2006) "Coding the Environment in Tactile Maps for Real-Time Guidance of the Visually Impaired". En IEEE International Symposium on MicroNanoMechanical and Human Science, pp. 1–6

Zhang, Z. (2000). A flexible new technique for camera calibration. Pattern Analysis and

Machine Intelligence, IEEE Transactions on, 22(11), 1330-1334.

