

Sistema de realidad virtual para exploración 3D con visión protésica simulada.

María Santos-Villafranca, Alejandro Pérez-Yus, Jesús Bermúdez-Cameo, José J. Guerrero

Afiliación: Grupo de Robótica, Percepción y Tiempo Real.
Instituto de Investigación en Ingeniería de Aragón (I3A)
Universidad de Zaragoza, Mariano Esquillor s/n, 50018, Zaragoza, Spain.
Tel. +34-976762707, e-mail: 727328@unizar.es

Resumen

Los recientes avances en las prótesis visuales han demostrado que es posible restaurar parte de la visión en determinados casos de deficiencia visual, aunque con limitaciones como la baja resolución (Figura 1) o un campo de visión reducido. Para experimentar más fácilmente se utilizan los simuladores de prótesis visuales (SPV).

Introducción

Se presenta un SPV inmersivo, flexible y realista mediante un sistema de realidad virtual, de forma que se puedan explorar entornos creados por un ordenador integrando el movimiento real de la cabeza del usuario en la simulación gracias a las gafas de realidad virtual Oculus Rift DK2 y permitiendo un desplazamiento por el escenario mediante un mando o joystick (Figura 2) para poder explorar el entorno de la forma más completa y realista posible.

SPV y Entorno virtual

Los puntos luminosos o fosfenos que ven las personas operadas, se han representado mediante círculos cuya iluminación es mas intensa en el centro que en los bordes y con un rango dinámico muy bajo, desde completamente apagados hasta completamente iluminados. Los fosfenos en la imagen se han dispuesto hexagonalmente ya que ofrece mayor densidad.

El nivel de intensidad de cada fosfeno depende de la representación de la imagen que se visualice (Figura 3), Downsampling, Modo de profundidad, Método de Canny, Método del suelo de ajedrez o el Método de detección de obstáculos y paredes.

El robot usado para modelar a la persona consta de una estructura montada sobre una base de Turtlebot 3 que integra sensores de odometría y la cámara Asus Xtion Pro Live, al que se le ha añadido un modelo de persona y se le ha colocado la cámara a la altura de

los ojos con los posibles movimientos de la cabeza implementados.

Se han exportado entornos de SweetHome 3D, ya que ofrece mas posibilidades que el desarrollo de escenarios de Gazebo, conserva las texturas y colisiones y dispone de mucha variedad de objetos realistas y detallados.

Sistema de realidad virtual

El sistema portátil de realidad virtual se conecta con el simulador de prótesis visuales a través de internet (socket TCP/IP), permitiendo que el procesamiento más intensivo se realice en un equipo servidor de sobremesa. Se utilizan dos equipos con diferentes sistemas operativos conectados por sockets con protocolo de internet TCP/IP, ya que es una forma rápida de intercambiar información sin estar físicamente próximos.

En el ordenador que sirve de interfaz con el sistema de realidad virtual, se ha implementado la parte del socket que realiza la labor de cliente. Tiene como finalidad enviar la información de la orientación de las gafas y recibir la trama de intensidades del mapa de fosfenos del servidor donde se ejecuta el entorno virtual.

Desde el servidor se recibe los ángulos de giro de las gafas y se comunican al entorno virtual, se genera en este equipo todo el procesamiento de la imagen para convertirla a fosfenos [1], [2] y se envía al cliente la luminosidad de cada uno.

Segmentación semántica

Se ha introducido en un nodo ROS la implementación en pytorch de modelos de segmentación semántica (Figura 4) con el dataset de escenas MIT ADE20K [5],[6]. Se usa la red neuronal pre-entrenada DilatedNet [3], [4], capaz de detectar de manera muy eficaz 150 objetos en total. Esta red

resulta de gran utilidad para resaltar los objetos mas importantes y facilitar la detección de estos en la visión fosfénica, además se puede ejecutar a tiempo real.

Conclusiones

Este SPV, además de trabajar en tiempo real, tiene la posibilidad de ser extrapolado, es decir, podría crearse un nodo que leyera la información de una cámara real en vez de usar la de Gazebo y poder seguir funcionando.

Es muy interesante en la realización de pruebas con usuarios para poder evaluar de forma efectiva la importancia de una exploración completa e inmersiva del entorno para reconocimiento de objetos, estancias y navegación por el mismo. Además de poder comparar el efecto de la segmentación de objetos a la hora de realizar dichas tareas.

A diferencia de otros simuladores, éste es capaz de representar diferentes entornos virtuales de interiores, con diferentes estancias, con la posibilidad de explorarlos completamente sin necesidad de desplazarse en la realidad.

Agradecimientos

Este trabajo ha sido financiado por el proyecto RTI2018-096903-B-I00 (MINECO/FEDER, UE)

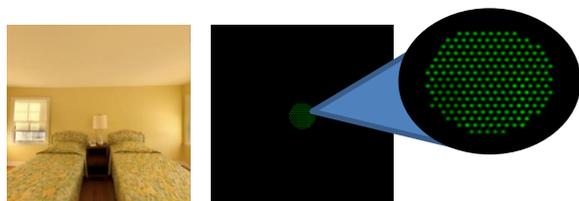


Figura 1. Diferencias entre la visión de una persona operada y otra con visión normal

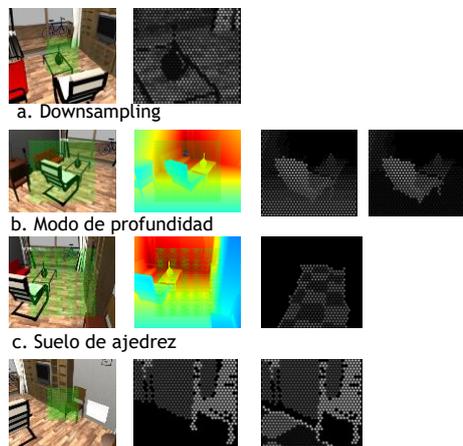


Figura 3. Representaciones de fosfenos

REFERENCIAS

- [1]. L. Montano Oliván, *Visual impaired guidance with simulated prosthetic vision in virtual environments*, Zaragoza. 2015.
- [2]. A. Perez-Yus, J. Bermudez-Cameo, G. Lopez-Nicolas, and J. J. Guerrero, *Depth and motion cues with phosphene patterns for prosthetic vision*, in IEEE International Conference on Computer Vision Workshops, 2017,
- [3]. L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, *Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs*, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834-848, 1 April 2018.
- [4]. Fisher Yu and Vladlen Koltun, *Multi-Scale Context Aggregation by Dilated Convolutions*, International Conference on Learning Representations (ICLR), May 2016.
- [5]. B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso and A. Torralba, *Scene Parsing through ADE20K Dataset*, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 5122-5130, doi: 10.1109/CVPR.2017.544.
- [6]. B. Zhou et al., *Semantic Understanding of Scenes through the ADE20K Dataset*. Accessed: Jun. 15, 2021.

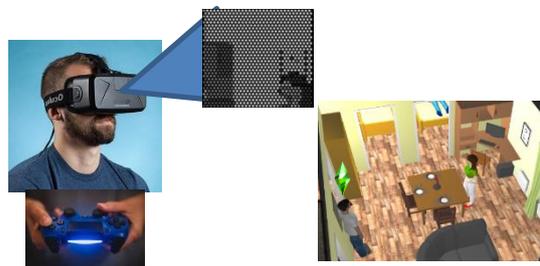


Figura 2. SPV en un entorno virtual con gafas de realidad virtual



Segmentación de objetos con una y tres pasadas de la red neuronal respectivamente



Figura 4. Resultado de la segmentación de objetos (arriba). Ejemplo de uso de la segmentación para destacar la figura de una persona con distintas representaciones de fosfenos (abajo)