



FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ciencias de la Computación
EL7008-1 Procesamiento Avanzado de Imágenes

Proyecto

Segmentación de objetos usando atención visual basada en
profundidad y movimiento

Integrantes: Camilo Jara D.
Profesor: Javier Ruiz del Solar.

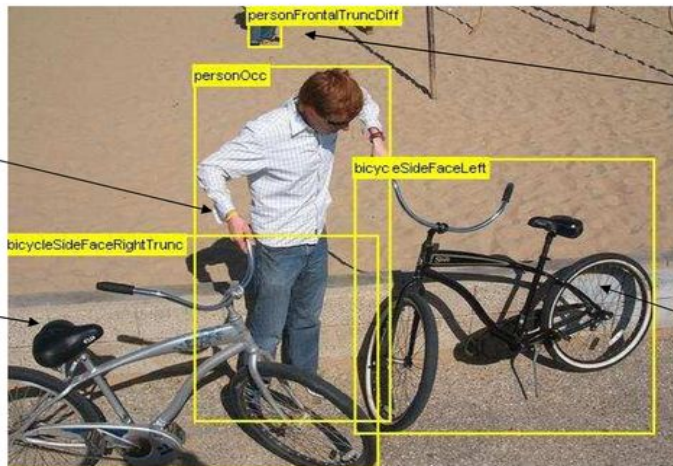
Introducción

Occluded

Object is significantly occluded within BB

Truncated

Object extends beyond BB



Difficult

Not scored in evaluation

Pose

Facing left



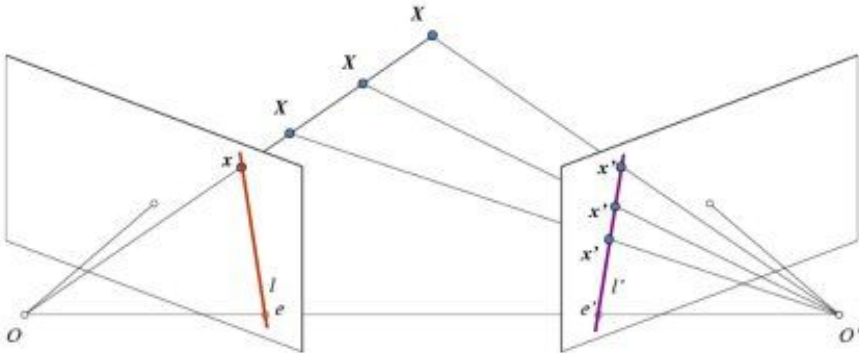
Objetivo

- Segmentar imágenes de manera robusta para ser entrenado posteriormente en clasificación de objetos.
- La secuencia de objetos será mostrada al robot Pepper por un “profesor”.
- Implementar método de segmentación mediante análisis de profundidad.
- Implementar método de segmentación mediante análisis de movimiento.

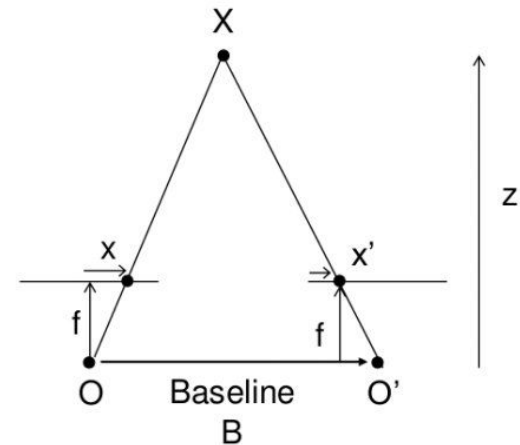
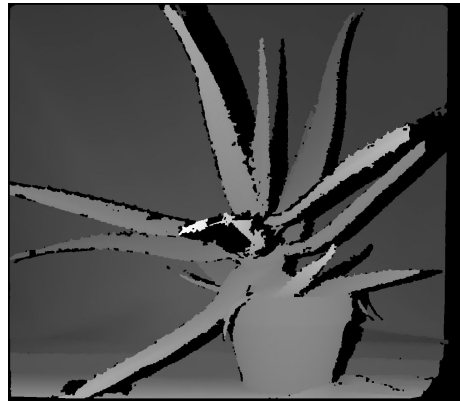
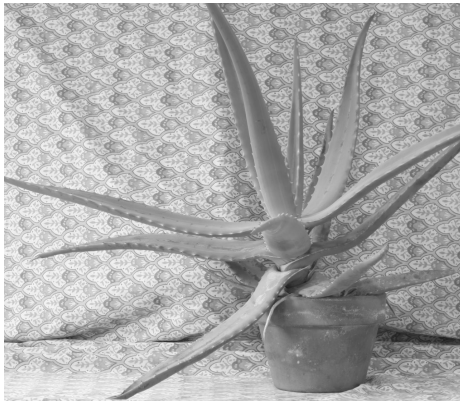


Marco Teórico

Análisis profundidad

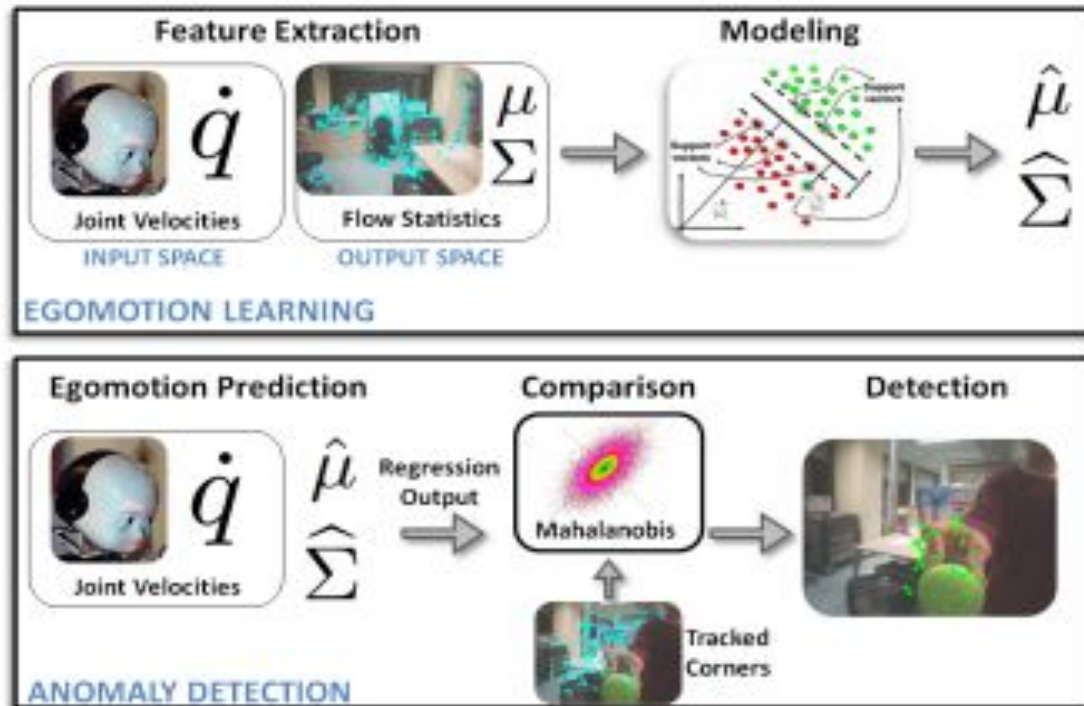


$$\text{disparity} = x - x' = \frac{Bf}{Z}$$



Marco Teórico

Análisis movimiento

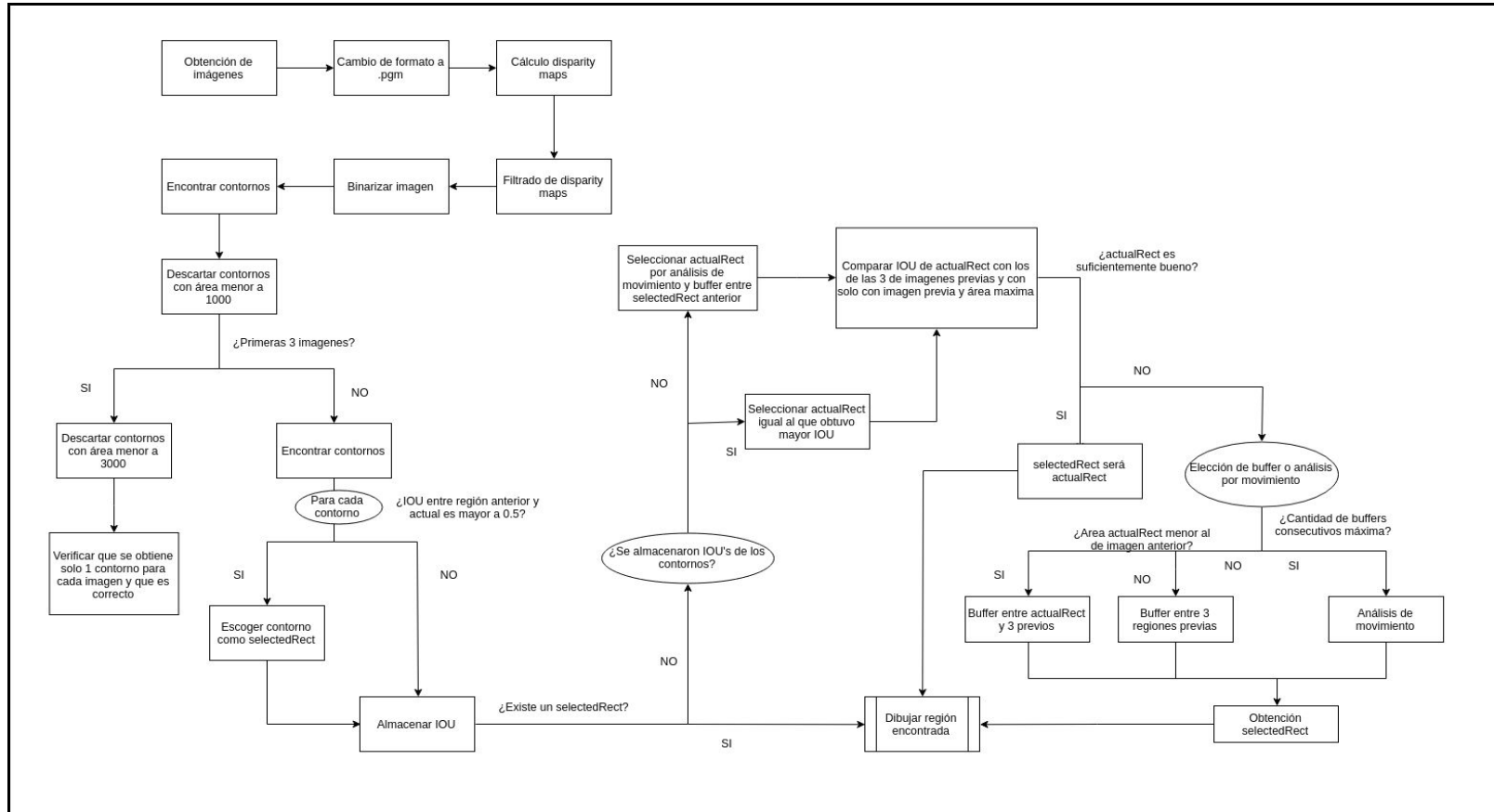


Dataset

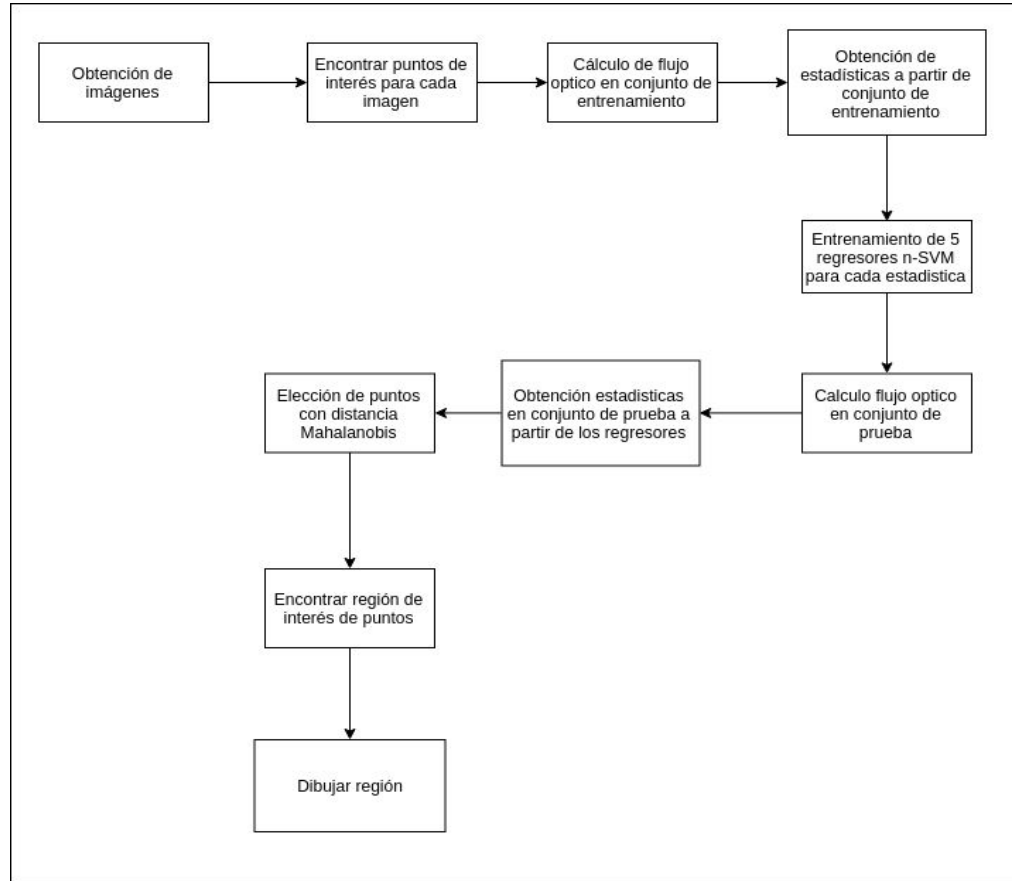
- ICub World Dataset [3] contiene imágenes obtenidas mediante una cámara RGBD ya rectificadas por lo que es directo calcular su disparity map.
- Posee 20 objetos distintos y 10 distintos para cada uno donde la dimensión de cada imagen es de 640x480.
- Para cada objeto se verifican distintos tipos de movimiento: rotación en 2 dimensiones, rotación en 3 dimensiones, modificación de escala, cambio de fondo (traslación) y un mix de todas.
- Además se realizan pruebas en distintos días a fin de tener distintos tipos de iluminación.



Segmentación por profundidad



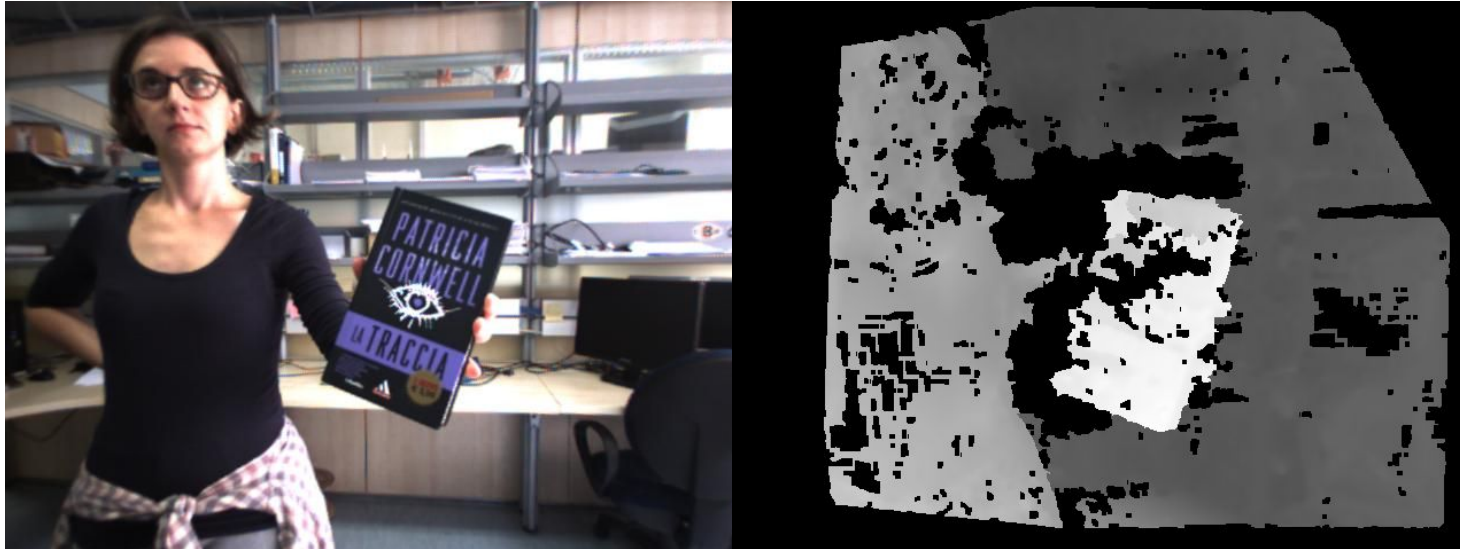
Segmentación por movimiento



Resultados

Análisis profundidad

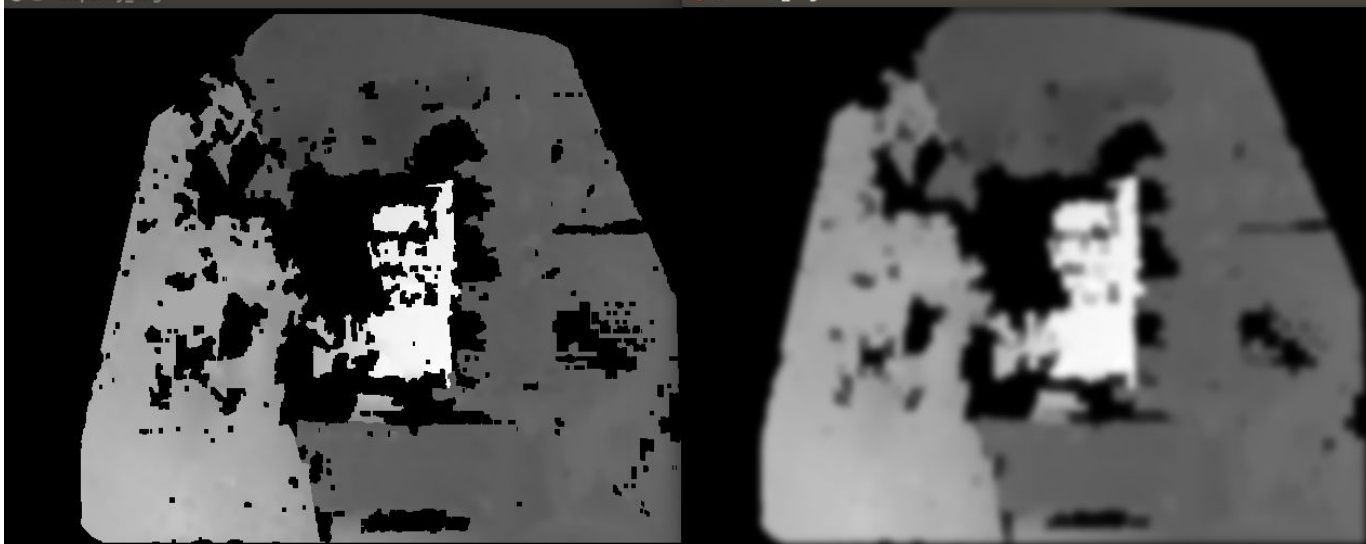
- Cálculo disparity map



Resultados

Análisis profundidad

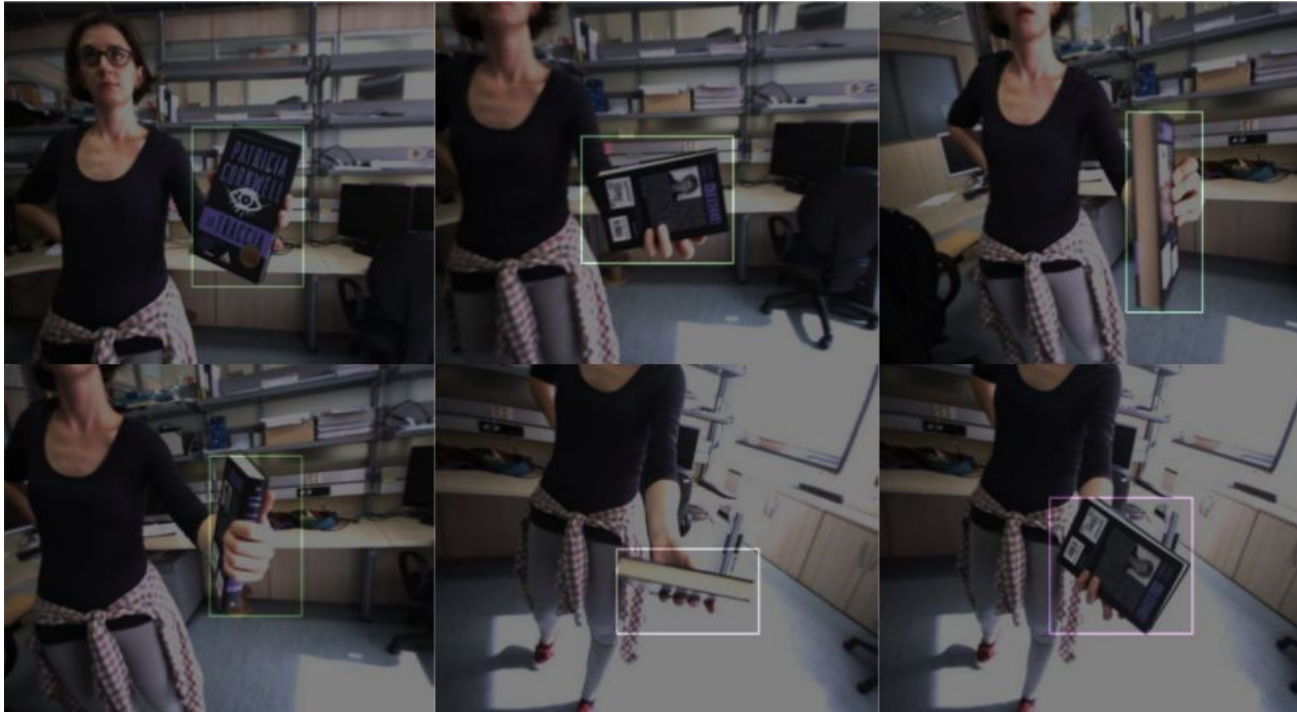
- Filtrado disparity map



Resultados

Análisis profundidad

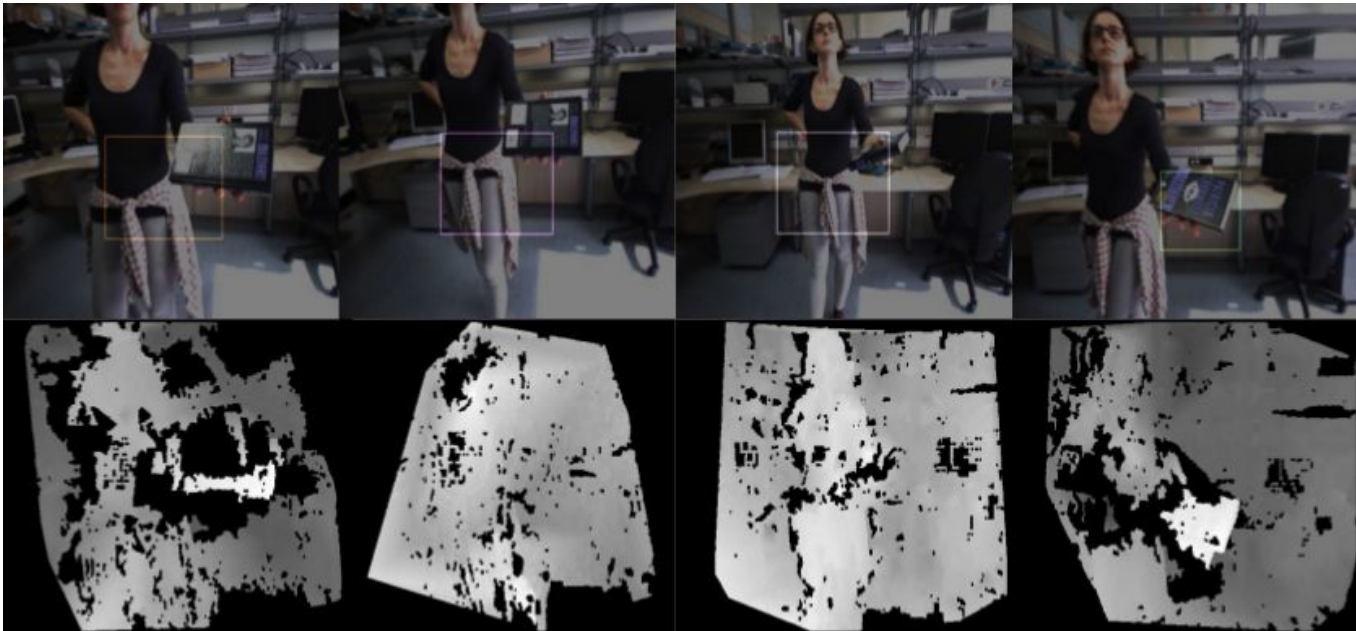
- Segmentación objetos



Resultados

Análisis profundidad

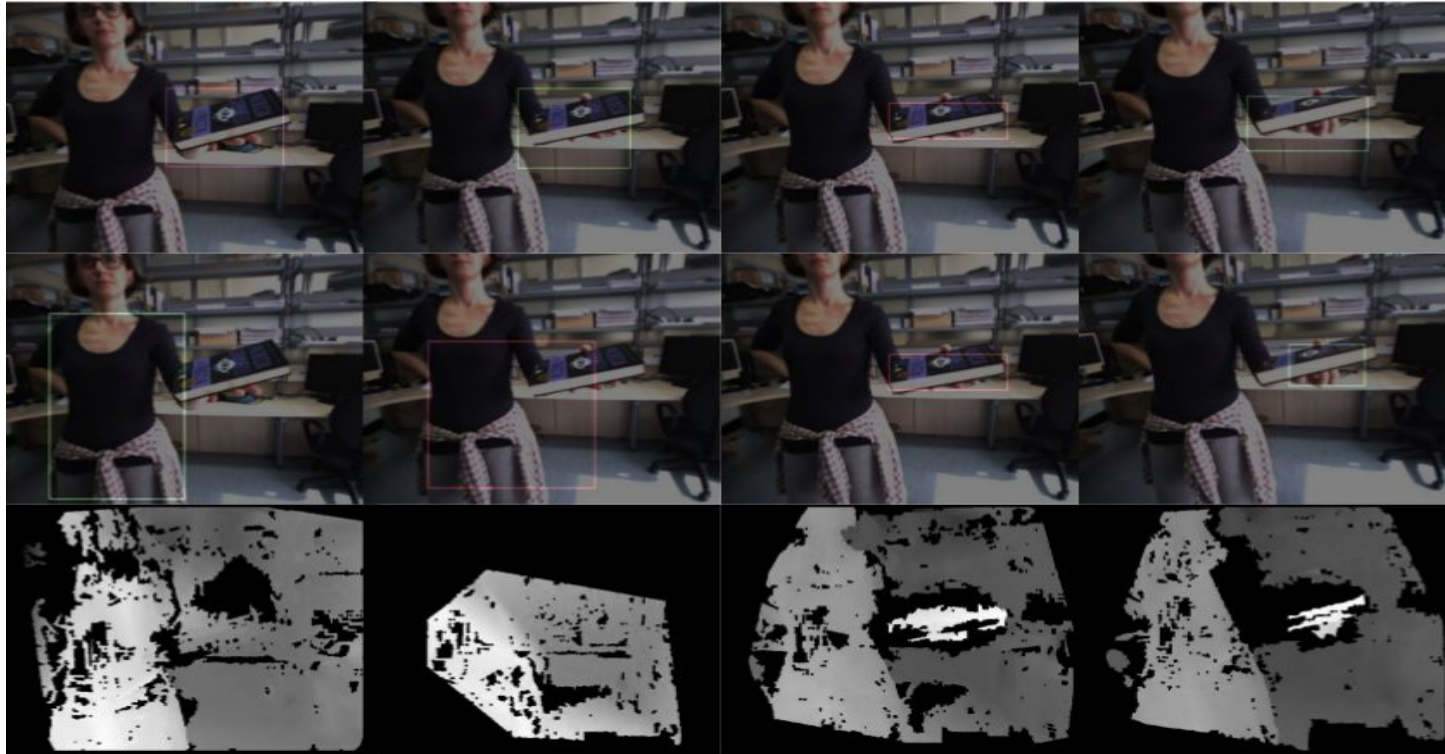
- Segmentación objetos (sin buffers)



Resultados

Análisis profundidad

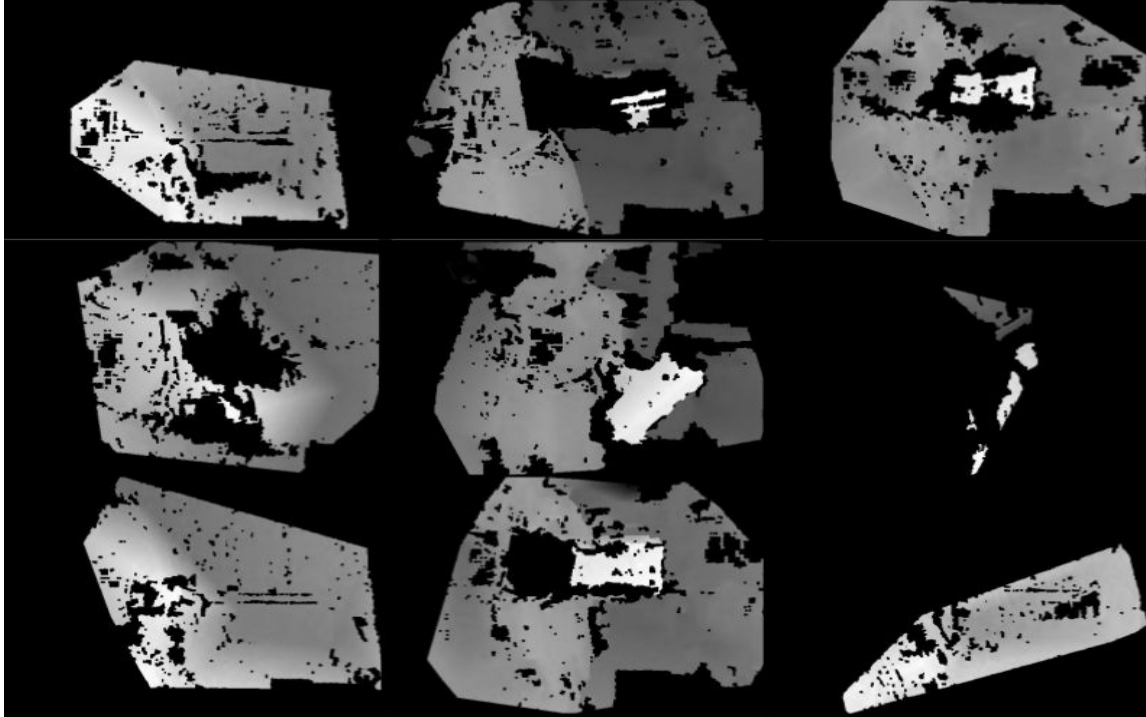
- Segmentación objetos (corrección con buffers)



Resultados

Análisis profundidad

- Segmentación objetos (eliminación de disparity maps)



Resultados

Análisis profundidad

- Segmentación objetos (modo de prueba)



Resultados

Análisis movimiento

- Segmentación objetos (obtención puntos interés)



Resultados

Análisis movimiento

- Segmentación objetos (libro)



Resultados

Análisis movimiento

- Segmentación objetos (celular)



Conclusiones

- El objetivo de la experiencia se logra tras segmentar correctamente las imágenes tomando las debidas restricciones.
- En análisis de profundidad las primeras 3 imágenes deben ser lo suficientemente buenas y por movimiento que el profesor tenga se mueva lento con respecto del objeto. Sin embargo, los métodos no son absolutos dada la dificultad del problema, para el método por profundidad se propone un modo de prueba a fin de ajustar debidamente los parámetros.
- Rápido procesamiento
- A fin de mejorar el método de RGB se propone utilizar background deletion y comparar ambos métodos. Un enfoque totalmente distinto es a partir de imágenes RGB aproximar las distancias con un enfoque de sampling no paramétrico usando SIFT Flow [6] o a partir de una red convolucional calcular el *disparity map* [7] con lo que a partir de ellas es posible utilizar el mismo enfoque de profundidad.

Referencias

- [1] Fanello, S. R., Ciliberto, C., Natale, L., & Metta, G. (2013, May). Weakly supervised strategies for natural object recognition in robotics. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on* (pp. 4223-4229). IEEE.
- [2] Pasquale, G., Mar, T., Ciliberto, C., Rosasco, L., & Natale, L. (2016). Enabling depth-driven visual attention on the iCub humanoid robot: instructions for use and new perspectives. *Frontiers in Robotics and AI*, 3, 35.
- [3] iCubWorld. (2018). Retrieved from <https://robotology.github.io/iCubWorld/>
- [4] Autonomous Vision Group | MPI for Intelligent Systems. (2018). Retrieved from <http://www.cvlibs.net/software/libelas/>
- [5] Satoshi Suzuki and others. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 30(1):32–46, 1985.
- [6] Karsch, K., Liu, C., & Kang, S. B. (2014). Depth transfer: Depth extraction from video using non-parametric sampling. *IEEE transactions on pattern analysis and machine intelligence*, 36(11), 2144-2158.
- [7] Laina, I., Rupprecht, C., Belagiannis, V., Tombari, F., & Navab, N. (2016, October). Deeper depth prediction with fully convolutional residual networks. In *3D Vision (3DV), 2016 Fourth International Conference on* (pp. 239-248). IEEE.