

Generación de nubes de puntos a través de un sistema de estereovisión para aplicaciones industriales y de investigación

PROBLEMA

El problema en la generación de nubes de puntos con sistemas de estereovisión radica en los desafíos técnicos como la calibración de cámaras, la elección del método correcto para que correspondencia de píxeles entre imágenes sea precisa y las inconsistencias en la escala de los datos 3D generados mediante los diferentes métodos.

OBJETIVO GENERAL

Desarrollar una solución tecnológica para la generación de nubes de puntos tridimensional mediante un sistema de estereovisión con dos cámaras analizando técnicas tradicionales y actuales para la obtención de mapas de disparidad, validando su desempeño a través de la medición de la altura de personas dentro del campo de visión del sistema.

PROPUESTA

La propuesta se centró en crear y analizar representaciones tridimensionales, conocidas como nubes de puntos 3D, utilizando la estereovisión. El proceso comenzó con la captura y calibración de la cámara. A partir de estas imágenes, se generaron mapas de disparidad, que reflejan las diferencias de posición de un mismo píxel en ambas imágenes. Para crear estos mapas, se aplicaron tres métodos: SGBM, un enfoque tradicional, y dos métodos basados en inteligencia artificial, RAFT-Stereo y Selective-IGEV. Luego, estas disparidades se transformaron en puntos en el espacio tridimensional, produciendo una nube de puntos detallada que permite visualizar la escena en 3D.



Obtención de datos

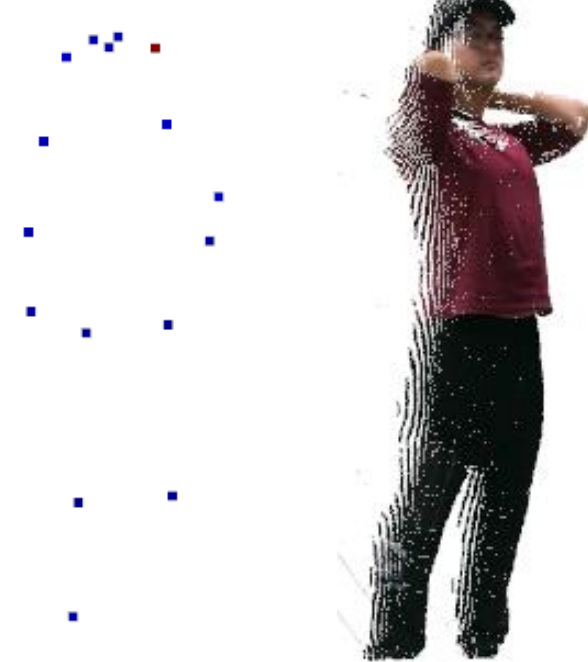
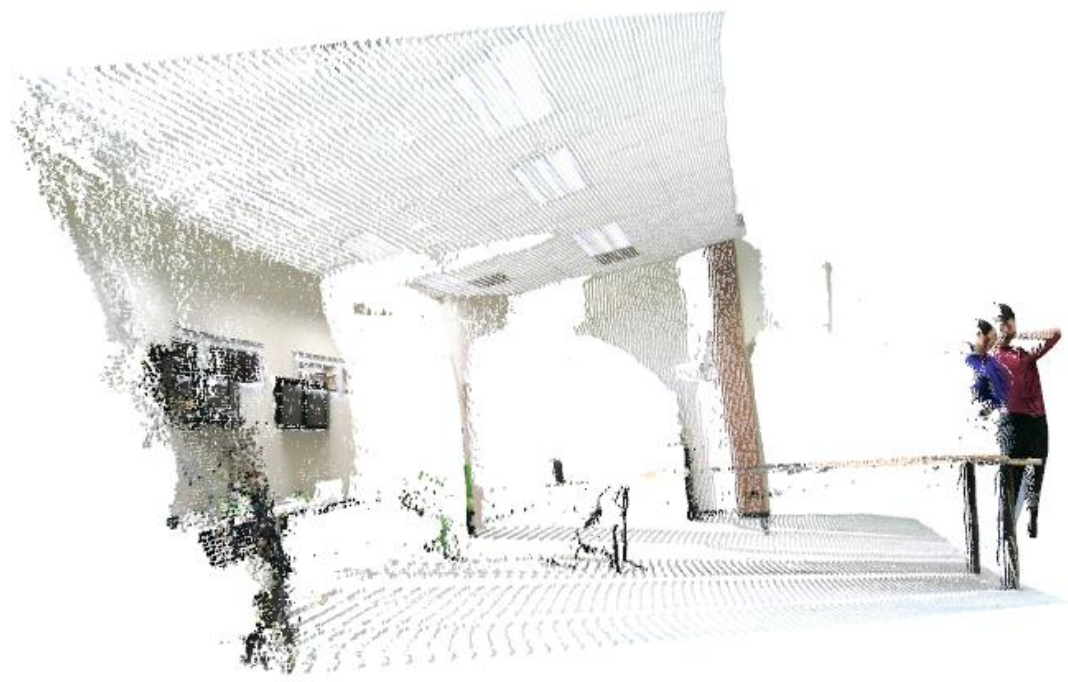
Calibración

Generación de mapa de disparidad

Generación de nube de puntos

Corrección

Cálculo de altura



Además, se crearon versiones menos detalladas de estas nubes de puntos para diversas aplicaciones. Se utilizó un modelo llamado YOLOv8 para identificar y segmentar objetos relevantes, como personas. Posteriormente, se aplicaron correcciones de escala y ajustes de profundidad. Finalmente, estas nubes de puntos simplificadas se emplearon para estimar la altura de las personas detectadas en diferentes escenas.

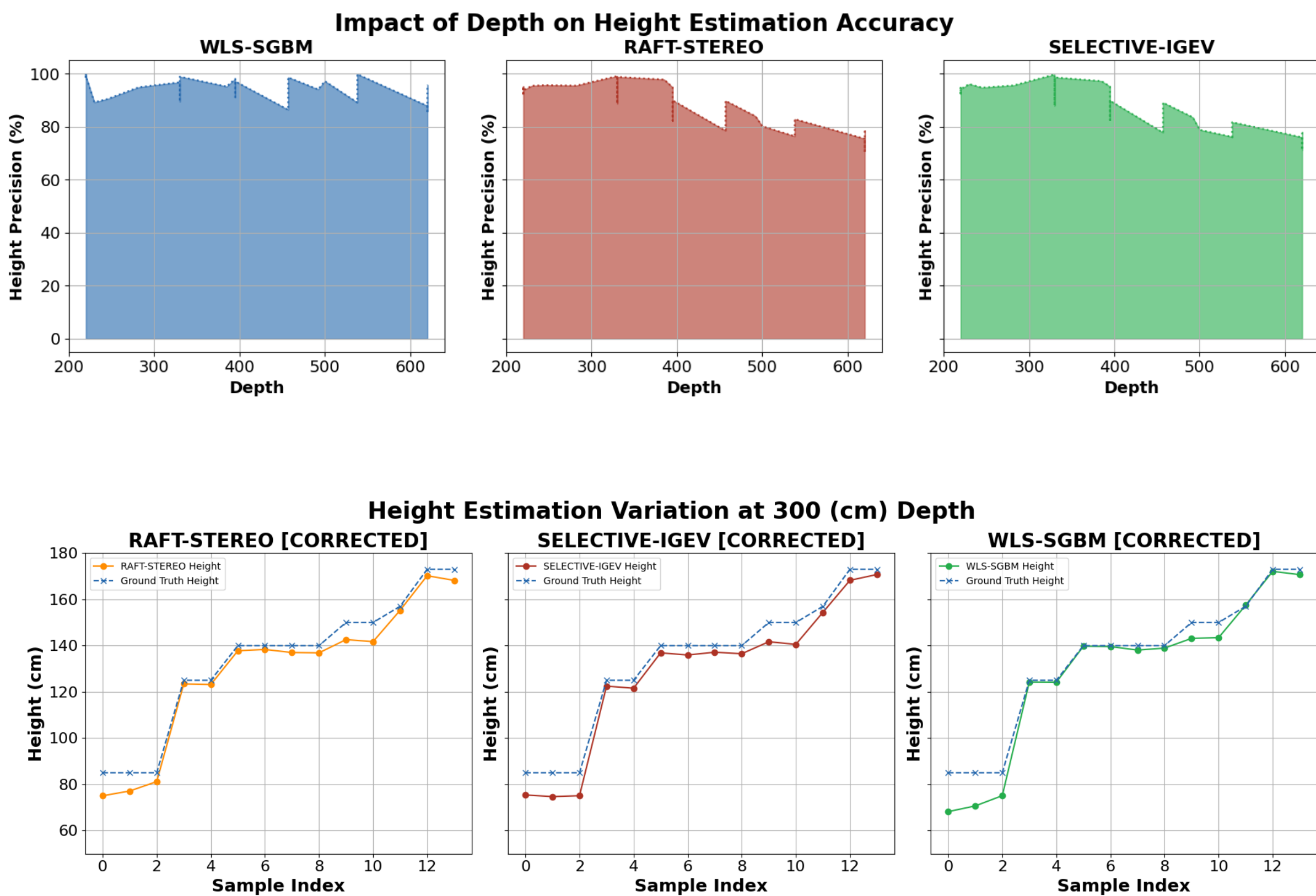
RESULTADOS

Los resultados muestran que la precisión en la estimación de profundidad y altura mejora significativamente tras la aplicación de correcciones a los modelos RAFT-Stereo, Selective-IGEV y WLS-SGBM.

Inicialmente, RAFT-Stereo y Selective-IGEV sobreestimaron la profundidad y altura, con errores de hasta 2000 y 400 unidades, respectivamente, por otro lado, WLS-SGBM hacia lo contrario, subestimaba ambos parámetros, pero con errores más consistentes.

Tras la corrección, todos los modelos se ajustaron a los valores reales, con estimaciones de profundidad entre 200 y 600 unidades y de altura entre 80 y 180 unidades, con desviaciones menores a 50 unidades.

WLS-SGBM mostró la mayor precisión promedio, con un 95.99% en la estimación de altura en profundidades entre 200 y 400 cm, superando a los modelos basados en machine learning.



CONCLUSIONES

- Los modelos generativos basados en ML están fuertemente influenciados por los sesgos presentes en sus *datasets* de entrenamiento, a diferencia de los métodos que emplean el cálculo directo de características específicas.
- RAFT-Stereo y Selective-IGEV generan mapas de disparidad de manera más homogénea, lo que los hace ideales para escenarios específicos donde se priorice reconstrucciones geométricas sobre la precisión.
- Modelos como RAFT-Stereo y Selective-IGEV pueden igualar la precisión de SGBM si se entrenan con *datasets* robustos en los rangos de profundidad necesarios.
- Métodos tradicionales como SGBM siguen siendo líderes en diversos escenarios, especialmente cuando se busca alta precisión a bajo costo.