Estimation of Human Hip Size Using ArUco Marker and Computer Vision

Leona Elizabeth Joseph (lj9984@rit.edu), Cecilia Alm (cecilia.o.alm@rit.edu)

RIT Rochester Institute of Technology

RIT Artificial Intelligence Master of Science Degree

INTRODUCTION

Goal:
The goal of this project is to estimate the hip size of a human using a 3D mesh, an ArUco marker with known dimensions, and an AI-driven approach.

Research Questions:

RQ1: How accurately can a computer vision model estimate the hip size of a human from images with varying angles and perspectives using an ArUco marker?

RQ2: How does the addition of different backgrounds (e.g., checkered, solid, complex) impact the accuracy of ArUco marker detection in varying environments?

METHOD

Dataset Creation:
The dataset includes 12 human body meshes that were imported into Unreal Engine, with a controlled environment. A camera trajectory was created to capture images automatically from different angles.

ArUco Marker Integration:
An ArUco marker with a known side length of 20 cm was placed next to the human meshes. The marker was used as a reference for estimating the hip size.

Figure 1: ArUco Marker example

Image Annotation:
63 images were captured for each mesh (total of 756 (63*12) images), and each image was annotated with ArUco and Hip as the two labels.

Model Training:
YOLOv8m was trained on the annotated dataset, and the detection accuracy was enhanced by introducing challenging backgrounds such as checkered patterns, ensuring better detection of ArUco markers.

Estimation:
After detecting the ArUco marker in the test set images, the horizontal distance of the hip was calculated using pixel values. The true horizontal distance of the ArUco marker (20 cm) was used to estimate the hip size.

DATASET

Figure 1 shows an example of the synthetic data. This image is a 3D mesh of a women with an ArUco marker and hip being detected.

Figure 2 shows the 12 3D meshes (6 women, 6 men) used to create the dataset. The meshes were created with different sizes using blender for dataset diversity.

Figure 2: An example image from the dataset

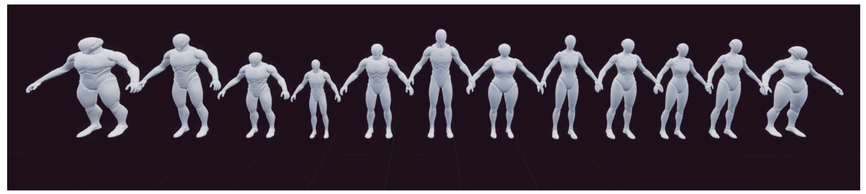

Figure 3: The different body types used for the dataset

RESULTS

- The ArUco and Hip detection has an overall precision of 0.982 and a recall of 0.986. The results are shown in more detail in **Table 1**.
- The Average Difference between the estimated Hip measurement and the true hip measurement is: 10.76 cm, Min Error: 5.85 cm, Max Error:16.93cm
- **Figure 3** shows the true hip measurements graphed with the estimated hip measurements

Class	Images	Count	Box(P)	R	mAP50	mAP50-95
All	70	140	0.982	0.986	0.981	0.749
ArUco	70	72	0.994	0.972	0.985	0.888
Hip	68	68	0.97	1	0.978	0.61

Table 1: Results for the ArUco and Hip detection

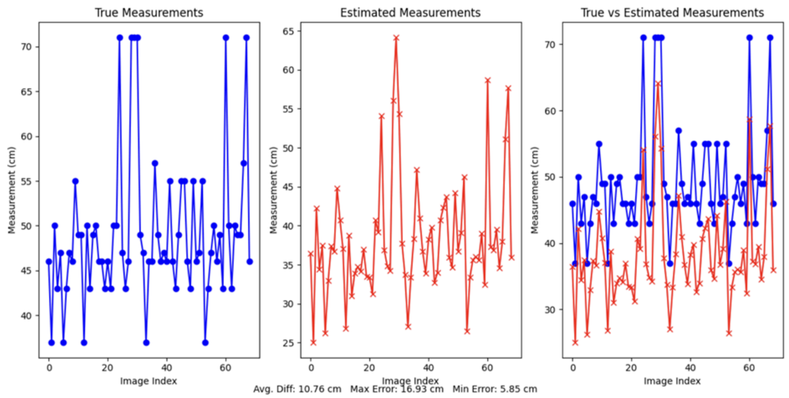

Figure 4: Comparison between the real and estimated hip sizes

LIMITATIONS

- The model is trained using synthetic data, and its performance may vary when applied to real-world images.
- The dataset only includes 12 human meshes (6 men, 6 women), limiting the diversity of body types.
- The current model does not account for slanted or angled images, which can distort the horizontal distance between the aruco marker and the hip.

DISCUSSION

This project explored the use of computer vision to estimate hip size using 3D meshes and ArUco markers. While the model demonstrated promising results, there is room for improvement in reducing estimation errors. A future goal can be the integration of depth estimation and testing againt real world data.

RQ1: The model achieved a mean difference of 10.7 cm in estimating hip size. This result is a good starting point. Methods to deal with slanted images, with perspective distortions will lead to better accuracy.

RQ2: Various unique environments helped the model reduce the number of false positives. The checkered backdrop resembles an ArUco marker making the model more robust to different test data.

REFERENCES

- Škorvánková, Dana, Adam Riečický, and Martin Madaras. "Automatic estimation of anthropometric human body measurements." arXiv preprint arXiv:2112.11992 (2021).
- Pollok, Thomas, et al. "UnrealGT: using unreal engine to generate ground truth datasets." Advances in Visual Computing: 14th International Symposium on Visual Computing, ISVC 2019, Lake Tahoe, NV, USA, October 7–9, 2019, Proceedings, Part I 14. Springer International Publishing, 2019.
- Ellis, Kenneth J. "Human body composition: in vivo methods." Physiological reviews (2000).