A Preliminary Evaluation of the Effect of Camera Motion on Infants' Movement Analysis based on RGB-D technology

D. Balta[1], I. G. Porco[2], T. Kanter[3], A. Cereatti[1], M. M. Schladen[4], P. S. Lum[4], U. Della Croce[2]
[1] Dept. of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
[2] Dept. of Biomedical Sciences, University of Sassari, Sassari, Italy
[3]Children's National Hospital, Washington, DC, USA
[4] Dept. of Biomedical Sciences, Catholic University of America, Washington DC, USA
Email: diletta.balta@polito.it

Summary

Markerless methods using pre-trained neural networks enable quantitative assessment of infant movements. This study evaluates the robustness of a method to compensate for camera motion, combining a single RGB-D camera, commercial software (*MediaPipe*) for 2D point tracking, and a custom algorithm for 3D coordinate estimation, occlusion handling, and camera motion compensation. To validate the approach, a doll was recorded from different camera angles, and its upper limb segment lengths estimates were compared to manual measures. Results show errors comparable to those obtained with a fixed camera and a manually trained network, supporting the method's clinical suitability.

Introduction

Markerless video analysis has emerged as a complement to clinical visual evaluation for quantitative infant movement assessment [1]. A key challenge is the need to move the camera to avoid occlusions while video recording during clinician evaluations. This study aims to evaluate the feasibility of a markerless method using an RGB-D camera, *MediaPipe* for 2D point tracking and a custom algorithm for 3D coordinate estimation, occlusion handling, and motion compensation [2,3].

Methods

An RGB-D camera (IntelRealSense D435, fs = 30 fps) was used for two video recordings of a doll, one with right arm straight and one with right arm bent, starting from a frontal view and then moving the camera through a range of views (Figure 1). *MediaPipe* was used for the 2D tracking of the following point of interest (PoIs): left and right shoulders (LS and RS), elbows (LE and RE), wrists (LW and RW), and hips (LH and RH). The centroid (B) coordinates of LH, RH, LS, and RS were calculated. PoI occlusions were dealt with as in [2]. The 3D coordinates of each PoI were estimated by combining 2D data with the relevant depth coordinate from depth images and referred to a local reference system centered in B [3] to compensate for the effect of camera movements. The upper arm (UA) segment length was defined as the 3D distance between the shoulder and the elbow PoIs, while the forearm (FA) segment length was defined as the 3D distance between the elbow and the wrist PoIs. These lengths were compared to reference values obtained with manual measures.

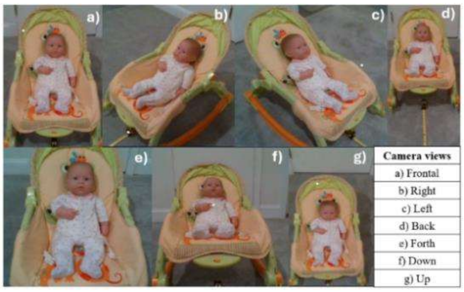

Figure 1: Different camera views included in the protocol.

Results and Discussion

The entire acquisition was segmented into intervals. Each interval included the movement of the camera from the frontal camera view to each specific target view. The Mean Absolute Error (MAE) of UA and FA length for each interval is presented in Table 1. MAE values are lower in the bent arm condition, mainly due to better wrist estimation and more accurate PoIs detection. As expected, the highest errors occurred when the arm was occluded by body parts (e.g. Right FA at Right view). In the absence of occlusions (frontal, back/forth and down/up views, see Figure 1), the MAE values were comparable to those obtained with the camera in a fixed frontal position and using a manually trained network [4].

Conclusions

Although a static frontal view remains the optimal condition for video recordings, this study proves the feasibility of acquisitions using a moving camera for clinical applications.

Acknowledgments

e.INS www.einsardinia.eu (Next Generation EU- PNRR - M4 C2 I1.5 CUP (J83C21000320007).).

References

[1] Silva et al. (2021). Research in dev. disabilities, 110
[2] Balta et al. (2022). Sensors, 19. GNB
[3] Balta et al. (2023). GNB Proceedings
[4] Balta et al. (2022) EMBC Proceedings.

Table I: MAE for right UA and FA segment's length at each camera view for both bent and straight conditions.

		Frontal	Right	Left	Back	Forth	Down	Up
Bent Arm	UA [mm]	3.9	5.0	3.8	4.0	2.6	3.7	4.6
	FA [mm]	4.4	21.4	3.9	7.8	5.9	4.4	7.9
Straight Arm	UA [mm]	14.2	16.9	7.8	12.6	12.9	18.0	13.9
	FA [mm]	13.7	15.2	10.4	17.9	17.7	16.7	18.5