Camera-based Analysis of Motion Coordination Between Infant Left and Right Limbs: A Clinical Study in NICU

Yiming Zhong[1], Ziyan Wu[1], Yongshen Zeng[1], Xiaoyan Song[2], Qiqiong Wang[2], Wenjin Wang[*]

***Abstract*—Limb movement coordination is a critical indicator in general movement analysis (GMA), which is often used to assess newborn neurological development. Asymmetry in limb movements may indicate brain injury or motor control disorders, also associated with conditions such as cerebral palsy. In this work, we present an automated video processing framework for assessing the coordination of left and right limb movements, aiming to assist healthcare professionals to evaluate infant's limb movement coordination during GMA. We use AggPose, a pose recognition tool based on a Transformer architecture, to extract 12 keypoints (including the arms and legs) from video frames. The intensity of movement is calculated using the temporal standard deviation of the keypoint coordinates. Finally, the coordination of movement is analyzed by comparing the cross-correlation and Pearson's correlation coefficients of the movement signals between left and right limbs. Our clinical dataset, created in the neonatal intensive care unit, includes 23 preterm infants without neurological disorders. The proposed method shows average cross-correlation and Pearson's correlation coefficients of 0.788 and 0.712, respectively, indicating the potential in analyzing the motion coordination of infant limb movements.**

Index Terms: Limb movement, Motion coordination analysis, Camera-based monitoring, Preterm infants, NICU.

I. Introduction

General Movement Analysis (GMA) is a tool for assessing neurological development based on movement patterns of infants [1], especially the infant's unstructured movements in a natural state such as stretching, bending, and rotating. GMA is primarily used for the early identification of infants who may have neurological developmental disorders, such as cerebral palsy [2]. An important indicator of GMA is the coordination of movements between left and right limbs [1]. In normal cases, the movement of an infant's left and right limbs is highly coordinated, meaning that the activities of the left arm and right arm, as well as the left leg and right leg, are synchronized. Whereas in abnormal cases, limb movements may appear asymmetrical or uncoordinated, particularly when there is clear ataxia, which could indicate underdevelopment of the brain and nervous system. However, GMA requires proficiency of caregivers in analyzing infant movement patterns. Therefore, our objective is to develop a camera-based monitoring system to automate the analysis of neonatal movement coordination, providing assistance to the neonatal intensive care unit (NICU) caregivers during GMA.

Wearable sensors have been used to capture spontaneous, general and fidgety movements in infants, producing promising classification results [3] [4]. However, these devices face several limitations. They offer only single-point measurements and lack of overview of different body-parts motions. Additionally, wearable devices can be uncomfortable, causing skin irritation and restricting infant movement, affecting their practical uses, especially in sensitive environments like NICU. Furthermore, the attachment and maintenance of these devices are time-consuming, also increasing the risk of infection among fragile infant patients.

To address the limitations of contact-based sensors, cameras have been exploited for monitoring vital signs, such as heart rate, heart rate variability, respiratory rate and SpO_2, in the NICU [5] [6] [7], offering a non-contact and more comfortable alternative. Previous research in Computer Vision enabled more comprehensive and detailed analysis of infant movements. Zeng *et al.* proposed an approach [8] that uses segmentation to generate masks for the upper limbs, lower limbs, torso, and head. Based on these masks, Pixflow [9] was applied to calculate the body-parts actigraphy. However, this approach has a limitation in the motion coordination assessment: segmentation is quite coarse, making it difficult to distinguish between left and right limbs for motion coordination analysis. For the task of analyzing the symmetry of infant body, we found that OpenPose [10] is commonly used for keypoint detection. The prior arts can be categorized into two groups. The first classifies infant postures as symmetrical or not based on a limited number of frames [11] [12]. The second calculates joint velocities and accelerations of the infant from the whole video sequence, extracts statistical features like the median and interquartile range to represent overall motion characteristics, and then compares left and right limb features to identify infant movement coordination [13] [14]. Although these methods rely on keypoint detection, they primarily focus on static posture classification or feature extraction, neglecting the motion cues in the time domain. This motivates us to explore this topic, as motion signal can

This work is supported by This work is supported by the National Key R&D Program of China (2022YFC2407800), National Natural Science Foundation of China (62271241, 62350068), Guangdong Basic and Applied Basic Research Foundation (2023A1515012983), Shenzhen Science and Technology Program (JSGGKQTD20221103174704003), Shenzhen Fundamental Research Program (JCYJ20220530112601003), and Shenzhen Medical Research Fund (D2402011).

[1]Department of Biomedical Engineering, Southern University of Science and Technology, China.

[2]Neonatal Intensive Care Unit, Nanfang Hospital of Southern Medical University, China.

[*]Corresponding author: Wenjin Wang (wangwj3@sustech.edu.cn).

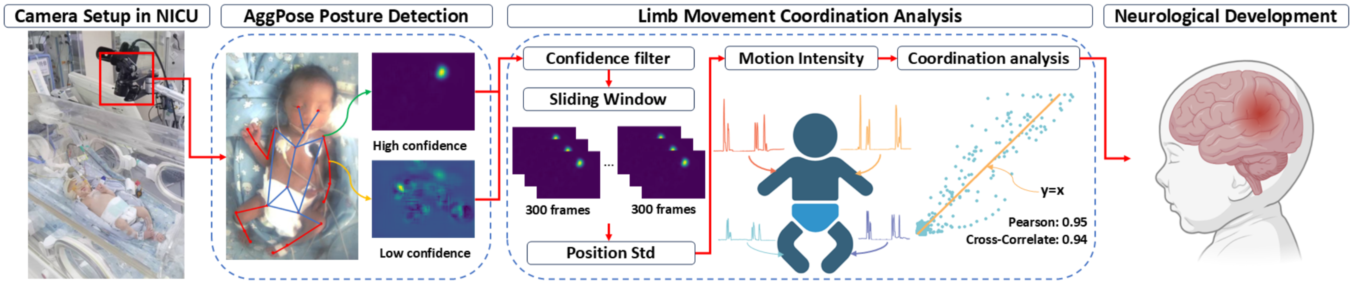


Fig. 1: Framework for analyzing neonatal limb movement and its neurological implication in NICU.

provide more direct information about infant limb movement coordination as compared to the static posture analysis.

We present a method that uses AggPose [15] to localize 21 keypoints and enable a more detailed analysis of their limb movements. Subsequently, the infant's limb movement intensity is calculated by using the keypoints' coordinates. Finally, we calculate cross-correlation and Pearson correlation coefficients between left and right limbs' movement intensities to quantify the motion coordination. To validate our approach, we conducted a clinical study on 23 newborns in NICU. By analyzing the coordination of their movements, we found that most of their cross-correlation coefficients (R-value) exceeded 0.6, exhibiting relatively high left-right limb coordination, which is consistent with their neurological development status. This also demonstrates that our method can assist caregivers in analyzing movement coordination, lowering the barrier for implementing GMA by reducing the need for extensive training or expert knowledge. It also supports the assessment of early neurological development in infants.

II. Materials and Method

A. Clinical Setup and Data Acquisition

Our clinical study was organized in the NICU of Nanfang Hospital of Southern Medical University, China (IRB no.: NFEC-2022-100). Under stable illumination conditions, the raw videos were recorded with an RGB camera (IDS-UI3860C, Germany) with a resolution of 484×274 pixels, sampled at 60 FPS. The camera was mounted directly above the infant, filming the entire body. Only infants in incubators with stable temperature conditions were recorded. Our dataset consisted of 23 infants (gestational age 28-40 weeks), each with a recording of at least 10 minutes in duration. Most of them were premature infants or suffered from respiratory distress. But no obvious neurological development problems were found.

B. Skeleton Landmark Detection and Motion Extraction

We use AggPose [15], a pose recognition model based on the transformer architecture, to compute the keypoint positions for each frame. Each frame is resized to 192×256 pixels for AggPose input, and for each keypoint, a 48×64 pixels heatmap is generated. This heatmap represents the model's confidence in detecting the keypoints, with the spatial coordinate of the maximum confidence to locate the joint's position.

Next, inspired by researches that use OpenPose [13] [14], we apply a sliding window of 300 frames (5 seconds) with a stride of 30 frames (0.5 second) to compute the average confidence for each keypoint within the window. Joints that do not surpass the confidence threshold are ignored and will not be used in the future analysis.

We then compute the standard deviation (std) of the positions for each sliding window (300 frames, 5s). The temporal std has two main advantages against the approach that takes the keypoints' coordinate differences between adjacent frames for estimating velocity and acceleration [13]. Firstly, calculating the variance of positions over a longer time period provides a stabler assessment of the movement intensity. This aligns better with GMA, which focuses on the overall movement patterns of the infant, such as fluidity, complexity, and variability. As a higher variance typically indicates more active and diverse movements, while lower variance may suggest reduced or rigid movement. Secondly, using std can reduce the influence of keypoint disturbances caused by the model (e.g., jitters), which will be encountered frequently in keypoint detection models. As a result, median and average smoothing are no longer necessary.

C. Coordination Analysis between Left and Right Limbs

To assess the coordination of limb movements, we conducted an analysis at two levels: keypoint signals and aggregated limb-level signals. For single keypoints, we selected joints such as the elbows, wrists, fingers, knees, ankles, and toes. At the limb level, we combined the movement signals from three keypoints (e.g., shoulder, elbow, and wrist for the arm; or hip, knee, and ankle for the leg) to generate an aggregated signal for each limb. That is : $M_{\mathrm{arm}} = M_{\mathrm{finger}} + M_{\mathrm{wrist}} + M_{\mathrm{elbow}}$ and $M_{\mathrm{leg}} = M_{\mathrm{toe}} + M_{\mathrm{ankle}} + M_{\mathrm{knee}}$.

During the actual recording process, the camera was not positioned perfectly above the head but had a slight tilt. This tilt caused an asymmetry in movement amplitudes of left and right limbs. However, we can reasonably assume that the actual lengths of left and right limbs (e.g., the left and right upper arms) are equal. By calculating the distance between the shoulder and elbow keypoints in video frames under the supine

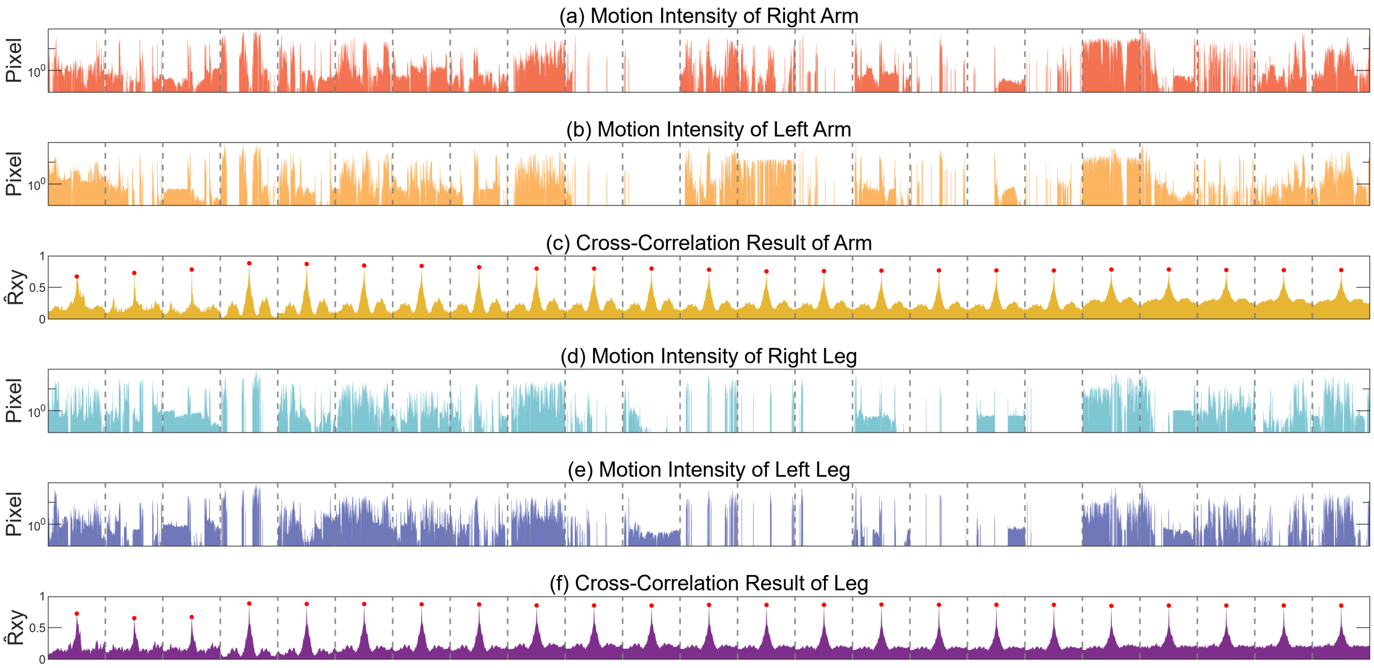


Fig. 2: The overall movement intensity of the arms and legs of 23 infants and their cross-correlation results.

posture, in other words, without any anti-gravity movements such as raising the arms, we can estimate the length of left and right upper arms based on joints detection. Subsequently, by normalizing the motion signals using the upper arm length, we can achieve a unified representation of the movement amplitudes for left and right limbs.

The motion signals of left and right limbs, are then compared with each other directly to calculate the cross-correlation and Pearson correlation coefficients. These metrics allow us to quantify the coordination of movements between the two sides, providing insight into the coordination and balance of limb activity.

The cross-correlation coefficient is defined as follows:

$$R_{xy}(m) = \sum_{n} x(n) \cdot y(n+m), \tag{1}$$

where m is the time delay (lag), indicating that the signal $y(n)$ is shifted by m time steps relative to itself.

$R_{xy}(m)$ is further normalized by:

$$\hat{R}_{xy}(m) = \frac{1}{\sqrt{R_{xx}(0) R_{yy}(0)}} R_{xy}(m), \tag{2}$$

where $R_{xx}(0)$ and $R_{yy}(0)$ represent the autocorrelation results of the x and y signals at zero lag, indicating the energy of the x and y signals. This step ensures that the autocorrelation results of the signals at zero lag will be 1.

It is worth noting that infant limb movements, particularly in newborns, often exhibit minor delays ranging from tens to hundreds of milliseconds [1]. So, after obtaining the cross-correlation results at different time delays, we search for the maximum cross-correlation coefficient within a reasonable time delay range (within 1 second), denoted as R_0, which represents the cross-correlation result. Additionally, we examine whether there are any cross-correlation coefficients at significant delays (exceeding 1 second) that are noticeably larger than R_0. If such cases exist, it would have implied a lack of synchronization between left and right limb movements, potentially signaling delayed neurological development.

Besides, the Pearson correlation coefficient is calculated as follows:

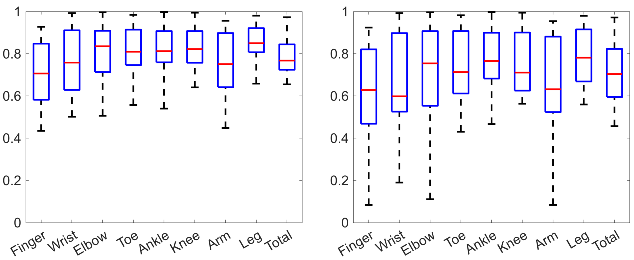
$$r = \frac{\sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{T} (x_t - \bar{x})^2 \sum_{t=1}^{T} (y_t - \bar{y})^2}}, \tag{3}$$

where x_t and y_t represent the movement signals of the left and right limbs at time t, and $\bar{x}$ and $\bar{y}$ denote their respective mean values. This method is more sensitive to the monotonic changes in the signals and reflects whether the movements on both sides occur simultaneously and with similar motion intensity changes.

By performing this coordination analysis, we can evaluate both the detailed movement characteristics at specific joints and the overall coordination of limb movements. The results of the cross-correlation and Pearson's coefficients are then used to quantify the infant's motion coordination.

III. Results and Discussion

For individual keypoints, the majority of cross-correlation maxima are within the reasonable delay range. However, in a few exceptional cases, certain keypoints in specific infants exhibit cross-correlation maxima with significant time shifts. These errors are mainly caused by the misidentification of keypoints and the disturbance of keypoints. For the coarse-grained analysis of arms and legs, all cross-correlation maxima are found with a time delay of zero, as shown in Fig.2.



(a) Cross-Correlation (b) Pearson Correlation

Fig. 3: Statistical analysis of motion coordination between left and right joints, as well as limbs.

For every single keypoint, most of the cross-correlation efficient exceed 0.7, indicating strong relationships. In detail, the average cross-correlation coefficients of finger, wrist, elbow, toe, ankle, knee, arm, and leg are 0.72, 0.76, 0.80, 0.80, 0.81, 0.83, 0.76, and 0.84, respectively, with an overall average of 0.79. Similarly, the corresponding Pearson correlation coefficients are 0.61, 0.67, 0.72, 0.73, 0.76, 0.76, 0.66, and 0.79, resulting in an overall mean of 0.71. Most of the correlation coefficients exceed 0.6, reflecting strong coordination in limb movements among these 23 infants. This aligns with the fact that studied 23 infants were in good neurological development condition.

Several improvements need to be made in the current study. First, the keypoint detection using AggPose exhibits some disturbance, affecting the calculation of limb's motion intensity, especially when the infant is still. In addition, during the recording process of some infants, our camera was not positioned directly above the infant, but rather with an angle. This may result in asymmetric motion amplitude between left and right limbs, potentially impacting the results. Furthermore, in some cases, the infants' arms and legs were obscured by clothing, making it difficult for AggPose to find the joints.

In conclusion, our findings reveal strong coordination in the limb movements of all the infants without neurological development issues, as indicated by the high correlation coefficients across keypoints. In the coarse-grained analysis of left and right limbs (arms and legs), the maximum values of their cross-correlation coefficient occur at the point without time delay. This highlights the precise synchronization of left and right limb movements occurring within a short time interval, which is crucial for evaluating infant's early motor and neurological development. These insights also provide a foundation for further studies on infant cognitive development and its potential applications in early detection of neurological disorders.

In the future, our method has potential for clinical applications in the assessment of infants suspected of neurological developmental disorders. By analyzing their movements, we aim to evaluate the sensitivity and specificity of this approach, providing a valuable tool for early diagnosis and intervention.

IV. Conclusions

In this work, we used AggPose to detect infant joints and temporal standard deviation to estimate their motion intensities. Cross-correlation and Pearson correlation coefficients are used to quantify the coordination of left and right joints and limbs motion. Our methods aligns better with the assessment on overall motion patterns during GMA. What's more, by comparing left and right limb movement intensity signals, our method captures signal details and temporal characteristics. The clinical evaluation on 23 newborns without neurological disorders in NICU shows that the measurements align well with their neurological development states. This approach has the potential to automate analysis of infant motor coordination between left and right limbs, supporting General Movement Analysis (GMA), and finally aiding in cerebral palsy assessment.

References

[1] Christa Einspieler and Heinz FR Prechtl. Prechtl's assessment of general movements: a diagnostic tool for the functional assessment of the young nervous system. *Mental retardation and developmental disabilities research reviews*, 11(1):61–67, 2005.

[2] Margot Bosanquet, Lisa Copeland, Robert Ware, and Roslyn Boyd. A systematic review of tests to predict cerebral palsy in young children. *Developmental Medicine & Child Neurology*, 55(5):418–426, 2013.

[3] Mohan Singh and Donald J Patterson. Involuntary gesture recognition for predicting cerebral palsy in high-risk infants. In *International Symposium on Wearable Computers (ISWC) 2010*, pages 1–8. IEEE, 2010.

[4] Dana Gravem, M Singh, et al. Assessment of infant movement with a compact wireless accelerometer system. *Journal of Medical Devices*, 2012.

[5] Wenjin Wang et al. Algorithmic principles of remote ppg. *IEEE Transactions on Biomedical Engineering*, 64(7):1479–1491, 2017.

[6] Yongshen Zeng, Chengyifeng Tan, et al. Camera-based monitoring of heart rate variability for preterm infants in neonatal intensive care unit. In *2023 IEEE International Conference on E-health Networking, Application & Services (Healthcom)*, pages 314–318. IEEE, 2023.

[7] Wenjin Wang and Albertus C den Brinker. Algorithmic insights of camera-based respiratory motion extraction. *Physiological measurement*, 43(7):075004, 2022.

[8] Y. Zeng et al. Video-based body parsing for neonatal body-parts actigraphy: A clinical study in nicu. In *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1–4, 2024.

[9] W. Wang and A. C. den Brinker. Algorithmic insights of camera-based respiratory motion extraction. *Physiological Measurement*, 43(7):075004, 2022.

[10] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[11] D. Ledwoń et al. Automated postural asymmetry assessment in infants neurodevelopmental evaluation using novel video-based features. *Computer Methods and Programs in Biomedicine*, 233:107455, 2023.

[12] I. Doroniewicz, D. J. Ledwoń, et al. Towards novel classification of infants' movement patterns supported by computerized video analysis. *Journal of NeuroEngineering and Rehabilitation*, 21(1):129, 2024.

[13] S. Ji, D. Ma, L. Pan, W. Wang, X. Peng, J. T. Amos, and P. Ren. Automated prediction of infant cognitive development risk by video: A pilot study. *IEEE Journal of Biomedical and Health Informatics*, 28(2):690–701, 2023.

[14] K. D. McCay, P. Hu, H. P. H. Shum, et al. A pose-based feature fusion and classification framework for the early prediction of cerebral palsy in infants. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:8–19, 2022.

[15] X. Cao et al. Aggpose: Deep aggregation vision transformer for infant pose estimation. *arXiv preprint*, arXiv:2205.05277, 2022.