SCALING DYNAMIC MODE DECOMPOSITION FOR REAL TIME ANALYSIS OF INFANT MOVEMENTS

Navya Annapareddy, Stephen Baek
School of Data Science
University of Virginia
Charlottesville, VA, USA
{na3au,baek}@virginia.edu

Lisa Letzkus, Santina Zanelli
Department of Pediatrics
University of Virginia
Charlottesville, VA, USA
{lmc8c,sz5d}@uvahealth.org

ABSTRACT

The analysis of characteristic motions in infants plays a pivotal role in quantifying developmental progress and clinical risk for neurodevelopmental and musculoskeletal abnormalities. Traditional methods often rely on resource intensive manual motion assessments carried out by clinicians while computer assisted approaches frequently utilize computationally expensive simulations or black-box classification models. These approaches struggle to efficiently both capture and differentiate the highly correlated dynamics of infant motion, limiting their ability to deliver actionable insights in a clinically viable decision time frame. In response to these challenges, we introduce the use of dynamic mode decomposition (DMD) as a transformative approach for decomposing complex infant motion into interpretable, independent components that are linearly additive in nature. DMD not only enables extraction of large scale clinically meaningful patterns but also can integrate with existing computer assisted interventions with regard to standardized motion features. We assess an optimized DMD formulation on 275,000 frames of infant motion in clinical settings that have undergone manual motion assessment by clinicians. Our experimental results show that using DMD modes as predictive components result in equal or superior accuracy in predicting abnormal clinical motion assessments compared to traditional manual or computer assisted methods. In addition, these modes serve as highly data informative features themselves and can be utilized as a novel basis for personalized clinical analysis and uncertainty quantification at scale.

1 INTRODUCTION

Infant motion analysis plays a critical role in developmental pediatric healthcare by extracting invaluable insights of cognitive and motor milestone trajectories and offering opportunities for early detection of potential abnormalities. By systematically tracking and analyzing infant movements, clinicians can identify deviations from normative developmental milestones and initiate timely interventions (Letzkus et al., 2024). However, the inherent complexity, variability, and dimensionality of motion systems pose significant computational barriers in extracting meaningful motion patterns and actionable insights in a clinically useful timescale. Infant movements are characterized by highly correlated, transient dynamics that are difficult to quantify using conventional methods. Notably, numerous clinical bodies of work have demonstrated infant movements correlated to high developmental disorder risk are characterized by a lack of fluency and independent extremity movements where bodily movements are restricted to highly correlated joints rather than the whole system.

Traditional methods to determine developmental progress based on predictive factors like movement variance and complexity often rely on manual assessments by individual clinicians and resource-intensive heuristic assessments with varying levels of inter-rater reliability (Crowle et al., 2017). Computer assisted interventions frequently feature classification models with little to no interpretable features or computationally expensive simulation frameworks and focus on the reconstruction of infant motion data with limited capabilities to extract clinically meaningful motion structures for analysis Schroeder et al. (2020). Such models are not well suited for real-time or near real-time analysis at scale, limiting their utility in clinical settings where low latency and repeated, longitudi-

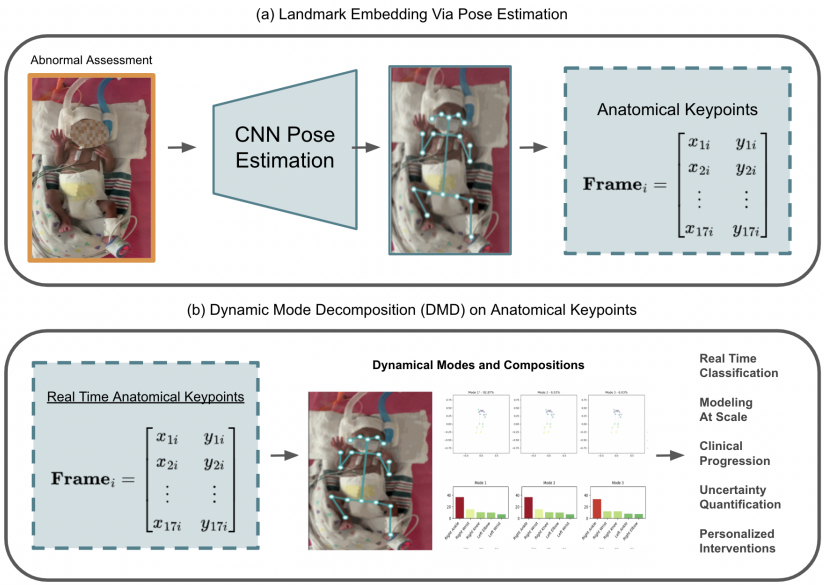

Figure 1: Pose estimation pipeline to derive anatomical keypoint features (a) for input into a dynamic mode decomposition (DMD) model (b) Top 3 most dominant DMD modes are shown with their most informational joints, as determined by their Frobenius Norm.

nal insights can inform interventions. There remains pressing need for analytical methodologies that can decompose these correlated motion dynamics into interpretable and independent components in a manner that preserves a low level of latency and is well suited for near real-time analysis and decision making.

Dynamic Mode Decomposition (DMD) offers a promising solution to this challenge. As a data-driven approach, DMD excels at identifying and decomposing complex systems into coherent modes that capture the underlying dynamics. By reframing infant motion as a complex dynamical system, DMD enables the extraction of meaningful patterns from highly correlated motion data, providing a computationally efficient and interpretable alternative to traditional simulations. This capability is particularly valuable in infant motion analysis where the ability to model transient dynamics with high fidelity is essential for clinical interpretation and decision-making. Moreover, DMD's efficiency in processing large datasets makes it well-suited for real-time applications, bridging the gap between theoretical models and practical, scalable solutions.

DMD is a well known analytical technique originating in physical systems that produces decomposition of a nonlinear system of states into independent components, even if the data in non-stationary in nature. While DMD originated in the study of physical systems, its novel application in the study of characteristic patterns in other dynamical systems such as motion detection and analysis has been proposed Erichson & Donovan (2016). Features extracted from high-dimensional motion data include characteristic poses, repetitive motions, and unique behavioral patterns that are essential for customizing rehabilitation and early intervention strategies based on individual motion profiles. DMD assumes that the dynamics of a physical system can be approximated by a linear combination of independent time relationships over time. Specifically, it assumes that a combination of modes, each with its own evolving exponential behaviors, can be decomposed from the larger system (H. Tu et al., 2014). For any given sequential snapshots X_1 and X_2 at different times, DMD learns a linear operator to map elements from one time step of a sequence of snapshots (state vectors) to the next. Typically, X_1 contains all snapshots except the last, and X_2 contains all snapshots except the first. Snapshots are data samples of a system at different times. For pose estimation problems, this means DMD consists of a linear approximations of non-linear dynamics over the observation window and motion that can be modeled as resulting from the sum, or superposition, of several independently interacting dynamic patterns.

While variance maximization methods like principal component analysis (PCA) and singular value decomposition (SVD) are effective for feature extraction and noise/dimensionality reduction, they do not account for any temporal relationships of the features in a given data feature set. DMD explicitly models the time dynamics of data and how data features evolve over time, and utilizes SVD to produce orthogonal matrices that provide a basis to project the original data, usually in a subspace smaller than the original.

2 METHODOLOGY

2.1 EVALUATION ON CLINICAL DATASET

Simple RGB videos of 275 active infants in supine position were acquired from an institutional neonatal intensive care unit (NICU) between July 2019 and November 2021. Motion was filmed in a standardized vertical orientation at a consistent angle above the infant at a frame rate of 30 frames per second (fps) and infants were free of positioning aids or motion restricting medical apparatus. Each video motion sequence was reviewed and scored by a minimum of 2 infant General Movement Assessment (GMA) certified clinician evaluators to be normal or abnormal. 17 standardized anatomical keypoints were extracted from each video frame in the motions store using the High-Resolution Network (HRNet) convolutional neural network (CNN). Unlike conventional CNNs that diminish resolution through pooling layers, HRNet initiates with a high-resolution sub-network and progressively integrates lower-resolution ones in parallel to capture detailed spatial information and accurately predict our infant joint landmarks with a mean average precision (mAP) of 94.2%. The center most 1000 frames of each infant's full motion store were concatenated as input into the optimized DMD model as the matrix X as representative motion sequences for each infant.

2.2 STANDARD DYNAMIC MODE DECOMPOSITION (DMD)

The standard formulation of DMD is utilized to analyze the dynamics from time series data represented as a matrix X, with each row capturing N features. This approach begins by dividing X into two matrices, X_1 and X_2, corresponding to sequential time frames. A reduced singular value decomposition (SVD) is then applied to X_1 targeting a specific rank r, formulated as $X_1 = U_r S_r V_r^*$. The output of this decomposition includes the matrices U, S, and V^*. The projected DMD operator A_{tilde} is computed as $A_{\text{tilde}} = U_r^T X_2 V_r S_r^{-1}$, and its eigenvalues and eigenvectors (W) are determined from the equation $A_{\text{tilde}} W = W\Lambda$. DMD modes Φ are then derived using $\Phi = U_r W$, and the state x_{k} at a given mode i and time k is reconstructed through $x_{\text{k}} = \Phi\Lambda^k \mathbf{b} = \sum_{i=1}^{r} \Lambda_i^k \mathbf{b}_i \phi_i$. Continuous time dynamics are defined by $x_{\text{k}} = \sum_{i=1}^{r} \mathbf{e}^{log(\lambda_i k)} \mathbf{b}_i \phi_i = \Phi \mathbf{e}^{\Omega k} \mathbf{b}$, with the return outputs being the DMD modes Φ, eigenvalues Λ, and vectors W.

Many variations of Standard DMD exist. Namely, Exact DMD extends the standard DMD formulation by incorporating the full utilization of the pseudo-inverse in the calculation of the DMD operator, augmenting DMD's ability to handle noisy or outlying data. This approach uses a similar segmentation of X into X_1 and X_2 but employs the full matrices U, S, and V^* from the SVD of X_1, leading to a DMD operator defined as $A = X_2 V S^{-1} U^*$. The application of the full pseudo-inverse typically results in a more accurate and robust approximation of the underlying dynamics, particularly useful in scenarios where the data set exhibits complex or more subtle dynamical structures (Brunton et al., 2016). Finally, we use the Hankel Time Delay embedding as a method enhance a given DMD system's observability Frame & Towne (2023). Time delay incorporates delayed copies of a given state into the analysis:

$$X_{\text{delayed}} = [X(t), X(t-1), \ldots, X(t-m+1)]$$

This approach extends the state vector to include m previous time steps (reported as a delay of d=m), effectively increasing the dimensionality of each snapshot by duplicating it (Curtis et al., 2023).

3 RESULTS AND DISCUSSION

To further minimize the role of environmental noise in DMD's application, we apply the bagging and optimization dynamic mode decomposition (BOPDMD) extension of the standard DMD framework.

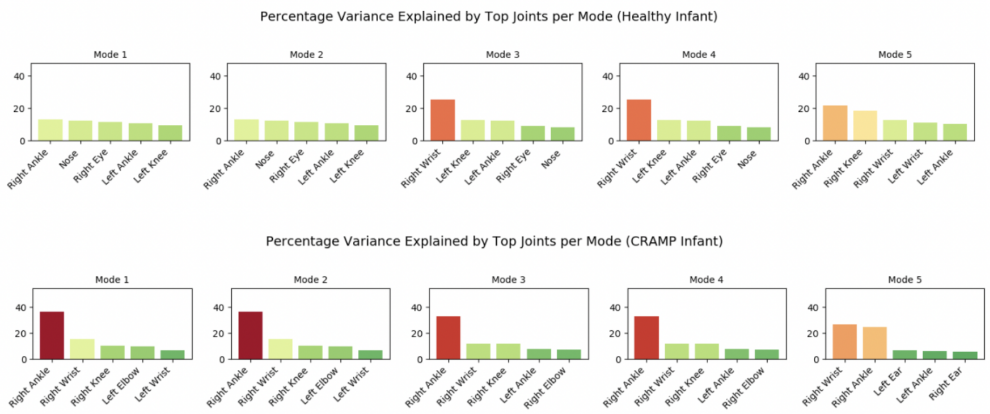

Figure 2: Joints that significantly contribute to the total variance of individual DMD modes for a health infant with equally variant joints (top) and an abnormal infant with biased joints (bottom) as coded by clinical infant movement assessment evaluators.

BOPDMD offers novel mechanisms for integrating statistical bagging, aggregation, subsampling, and ensembling techniques to improve DMD stability and accuracy. Given a dataset X and a number b of bootstrap samples, BOPDMD operates by generating b bootstrap samples X_i from X. Standard DMD is then applied independently to each bootstrap sample. The resulting dynamic modes and their corresponding eigenvalues from each sample are optimized and aggregated by taking the mean of the weights Ω_i, dynamic modes Φ_i, and temporal coefficients b_i from each sample. Additionally, regularization techniques are used to optimize each set of Ω_i, Φ_i, and b_i before aggregation. The optimized modes are aggregated from all bootstrapped samples to form a consensus set of dynamic modes. The BOPDMD method not only enhances the robustness of the decomposition by reducing the variance of the dynamical mode estimates but also can parameterize the dynamic modes to be more representative of the true underlying system dynamics (Sas).

Comparing the results of Exact DMD with and without Hankel delay, optDMD, and BOPDMD yields the most superior results for the optDMD formulation. Notably, the results for an ill configured time delay have higher error than Exact DMD alone and optDMD performs better than BOPDMD, its equivalent formulation with bagging, indicating potential for a more tolerant ensembling scheme. Finally, we utilize the makeups of the dynamical modes as a classification feature by taking the percentage variance explained by the top-10 most informational joints for a given mode. Even with such a severely restricted feature set, a simple support vector machine (SVM) with a radial basis function (RBF) is able to classify the clinician labeled clinical targets with a 95.2 % agreement rate with clinician diagnosis, equivalent to or exceeding the established baseline positive identification rates of both the GMA and deep learning based computer assisted interventions (Garfinkle, 2024; Gong et al., 2022; Reich et al., 2021).

4 Conclusion

Dynamic Mode Decomposition (DMD) emerges as a novel method to leverage the vast amounts of near-real time data available from feature extraction methods like pose estimation in infant motion analysis. Not only can DMD serve as a method of identifying large scale characteristic motions of infants, it can serve as a form of feature and dimensionality reduction via its dynamical modes. Unlike quantitative methods that cannot decouple motion characteristics to the temporal dimension, DMD modes are independent to each other, successfully decouple spatial and temporal dynamics, and are linearly additive. The properties make DMD components extremely interpretable, as they enable analysis, reconstruction, and manipulation of each dynamic component and their representative features. Moving forward, future work will focus on refining the DMD ensembling techniques, expanding the feature set to build more robust model ablations across clinical targets, and integrating

these methods directly into real-time clinical workflows via real-time streaming or classification to further validate the utility of analytical methods at scale in infant motion analysis.

REFERENCES

Bagging, optimized dynamic mode decomposition for robust, stable forecasting with spatial and temporal uncertainty quantification. 380. ISSN 1471-2962.

Bingni W. Brunton, Lise A. Johnson, Jeffrey G. Ojemann, and J. Nathan Kutz. Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods*, 258:1–15, January 2016. ISSN 0165-0270. doi: 10.1016/j.jneumeth.2015.10.010.

Cathryn Crowle, Claire Galea, Catherine Morgan, Iona Novak, Karen Walker, and Nadia Badawi. Inter-observer agreement of the general movements assessment with infants following surgery. *Early human development*, 104:17–21, 2017.

Christopher W Curtis, D Jay Alford-Lago, Erik Bollt, and Andrew Tuma. Machine learning enhanced hankel dynamic-mode decomposition. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33(8), 2023.

N Benjamin Erichson and Carl Donovan. Randomized low-rank dynamic mode decomposition for motion detection. *Computer Vision and Image Understanding*, 146:40–50, 2016.

Peter Frame and Aaron Towne. Space-time pod and the hankel matrix. *Plos one*, 18(8):e0289637, 2023.

Jarred Garfinkle. Automation to approximate the gestalt: applying machine learning to the general movement assessment. *Pediatric Research*, pp. 1–2, 2024.

Xiaohui Gong, Xiao Li, Li Ma, Weilin Tong, Fangyu Shi, Menghan Hu, Xiao-Ping Zhang, Guangjun Yu, and Cheng Yang. Preterm infant general movements assessment via representation learning. *Displays*, 75:102308, 2022.

Jonathan H. Tu, Clarence W. Rowley, Dirk M. Luchtenburg, Steven L. Brunton, and J. Nathan Kutz. On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1(2):391–421, 2014. ISSN 2158-2505. doi: 10.3934/jcd.2014.1.391.

Lisa Letzkus, J Vince Pulido, Abiodun Adeyemo, Stephen Baek, and Santina Zanelli. Machine learning approaches to evaluate infants' general movements in the writhing stage—a pilot study. *Scientific Reports*, 14(1):4522, 2024.

Simon Reich, Dajie Zhang, Tomas Kulvicius, Sven Bölte, Karin Nielsen-Saines, Florian B Pokorny, Robert Peharz, Luise Poustka, Florentin Wörgötter, Christa Einspieler, et al. Novel ai driven approach to classify infant motor functions. *Scientific Reports*, 11(1):9888, 2021.

A Sebastian Schroeder, Nikolas Hesse, Raphael Weinberger, Uta Tacke, Lucia Gerstl, Anne Hilgendorff, Florian Heinen, Michael Arens, Linze J Dijkstra, Sergi Pujades Rocamora, et al. General movement assessment from videos of computed 3d infant body models is equally effective compared to conventional rgb video rating. *Early human development*, 144:104967, 2020.