

WHOSE MOVEMENT CAN BE MEASURED, WHOSE MUST BE AUTHORED BY HAND, AND WHAT IS LOST IN THE GAP?

STATEMENT OF RESEARCH PROJECT

Motion capture and pose estimation now underwrite a great deal of how human movement is recorded, analysed and reproduced — in clinical assessment, in sports science, in animation and virtual production, and in the training data of generative motion models. These systems are presented as neutral instruments. They are not. Like face recognition before them (Buolamwini & Gebru, 2018), they are trained on a narrow distribution of bodies performing a narrow repertoire of movements, and they fail unevenly. Preliminary work by the principal investigator shows that per-joint reconstruction error in egocentric body recovery varies systematically with body-shape parameters [Fish2Mesh citation], which means the instrument encodes an implicit norm of what a body is.

This project asks what that norm excludes, and what it costs. It does so by pairing computational measurement with animation practice, because animators have been solving the same problem by hand for a century. Where a capture system fails to register a movement, an animator must decide what that movement means and construct it pose by pose. Animation is therefore not merely an application domain here; it is a second, older instrument for rendering movement legible, with its own assumptions, its own craft vocabulary, and — crucially — an explicit author who can be asked why.

The central research question is: whose movement can be measured, whose must be authored by hand, and what is lost in the gap?

Three sub-questions follow. First, across which bodies and which kinds of movement do current motion estimation systems fail, and are those failures random or structured? Second, what do animators do when they represent movement that cannot be captured — grief, hesitation, ritual, the bearing of someone under duress — and what knowledge is embedded in those decisions? Third, what would a movement corpus and evaluation protocol look like if it were built from the movements these systems currently miss, rather than from the movements they already handle well?

The project is deliberately sited in Asia and in movement traditions that are poorly represented in existing motion datasets, which are overwhelmingly Euro-American, laboratory-recorded, and biased toward locomotion and sport. Gesture, ritual, mourning practice, devotional movement and the physical vocabularies of displacement are largely absent from the corpora on which these systems are trained and evaluated.

Why this pairing, specifically

The co-investigating perspective from the School of Art, Design and Media is not consultative. Determining what a movement means — whether a hesitation is fatigue or fear, whether a gesture is culturally specific or idiosyncratic, whether a reconstruction has preserved the thing that mattered — is not derivable from reconstruction error. It is an authorial judgement, and it is the daily work of a director who builds performance frame by frame. Stop-motion practice is particularly apt: because every pose is deliberately placed, the animator can articulate the reasoning behind movement in a way that neither motion capture subjects nor generative models can.

References

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77–91.

SCOPE OF WORK FOR SELECTED PHD STUDENT

The student will produce a characterisation of where human motion estimation fails, an account of how animation practice compensates, and a corpus and protocol built from the gap between them.

Year 1 establishes the technical baseline. The student surveys motion estimation methods and the corpora they are trained and evaluated on, documenting the demographic and movement-type composition of those corpora — which is itself a publishable contribution, since it is rarely stated. They then reproduce and extend existing body-conditioned error analysis, moving from body shape alone to movement category, clothing and occlusion.

Year 2 runs in parallel with the practice-led track. The student works alongside ADM animators, collecting production material and conducting stimulated-recall sessions. This requires the student to become literate in animation practice rather than merely to observe it, and the project expects them to complete foundational animation training within the School. Concurrently they design the corpus: what movements, performed by whom, recorded how, and annotated against what scheme.

Year 3 collects and annotates the corpus and runs the evaluation, comparing how current systems perform on it against their reported benchmark performance. Annotation is conducted with practitioners qualified to judge the movements in question, with inter-annotator agreement reported.

Year 4 completes the analysis and the thesis, and delivers the corpus and protocol in reusable form.

The student profile is deliberately unusual: someone comfortable with quantitative evaluation who is also willing to spend a year in an animation studio. Candidates may come from computer vision, communication, media studies, or animation practice with quantitative aptitude. The project is well suited to a candidate whose background does not sit neatly in one discipline.

Ethical and practical constraints — consent for recorded movement, the cultural sensitivity of ritual and mourning practice, ownership of contributed material, and the terms on which a corpus of culturally specific movement is released — are treated as a substantive component of the work rather than as compliance, and will be addressed with the communities from whom movement is recorded.