



Identity Beyond Appearance: Reversing Human Digital Twins

Sirisha Talapuru, Ram Dantu, Vinh Quach, Shakila Zaman, Apurba Pokharel

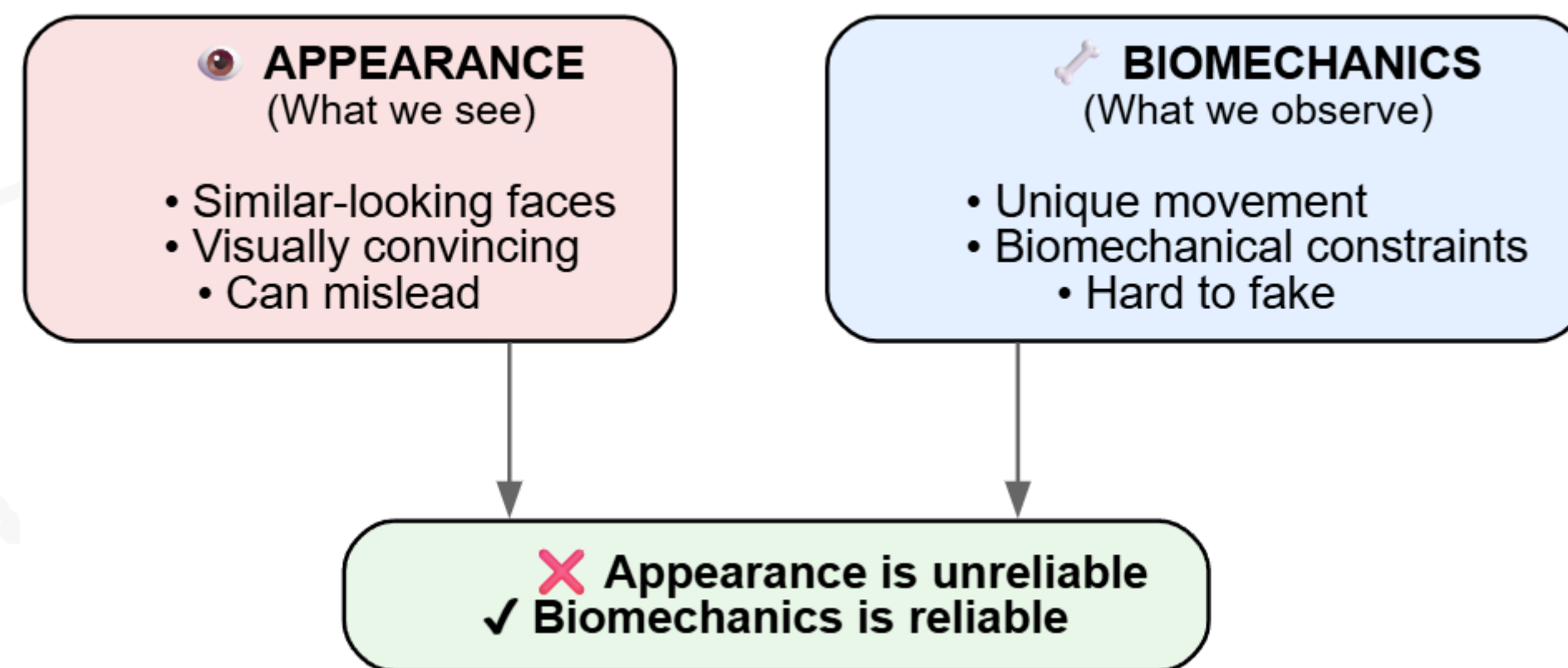
University of North Texas, Denton, TX



Appearance Does Not Define Identity

- Visual evidence is often used to identify individuals, but it can be misleading. In a recent incident at Brown University, surveillance footage led to the identification of the wrong person, highlighting the limits of appearance-based analysis.
- In this work, we adopt a **Reverse Human Digital Twin (HDT) framework** to verify identity using consistency in human behavior and texture. Given a trusted HDT and an observed video, we analyze biomechanical patterns and temporal texture variations and compare them against the HDT to determine whether the observed behavior aligns with the true individual or indicates impersonation.

Appearance vs Biomechanics



Limitations of Artifact-Based Detection



Relies on visual artifacts such as texture and blending errors



Artifacts disappear after compression and re-encoding



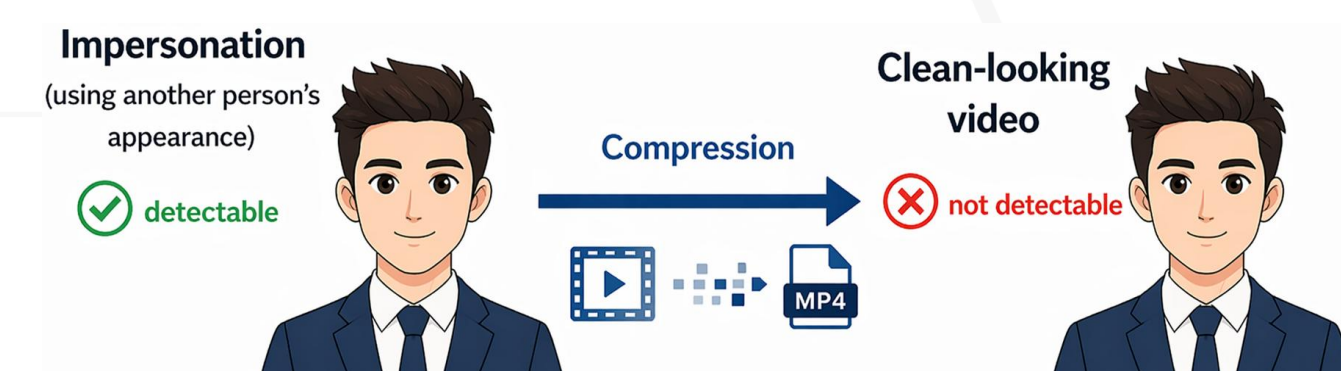
Limited generalization to unseen manipulation techniques



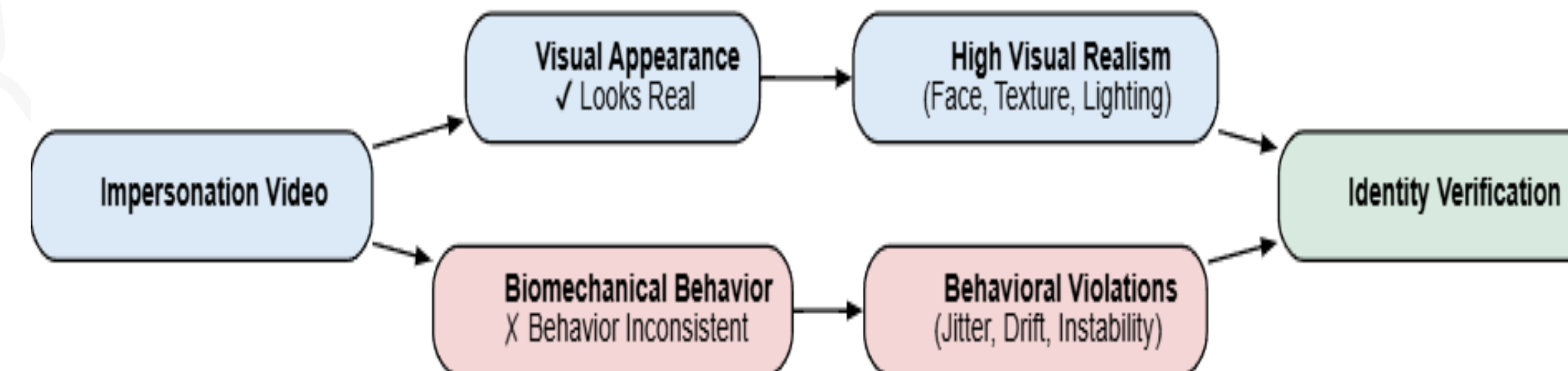
Sensitive to lighting and video quality variations



Ignores underlying biomechanical inconsistencies



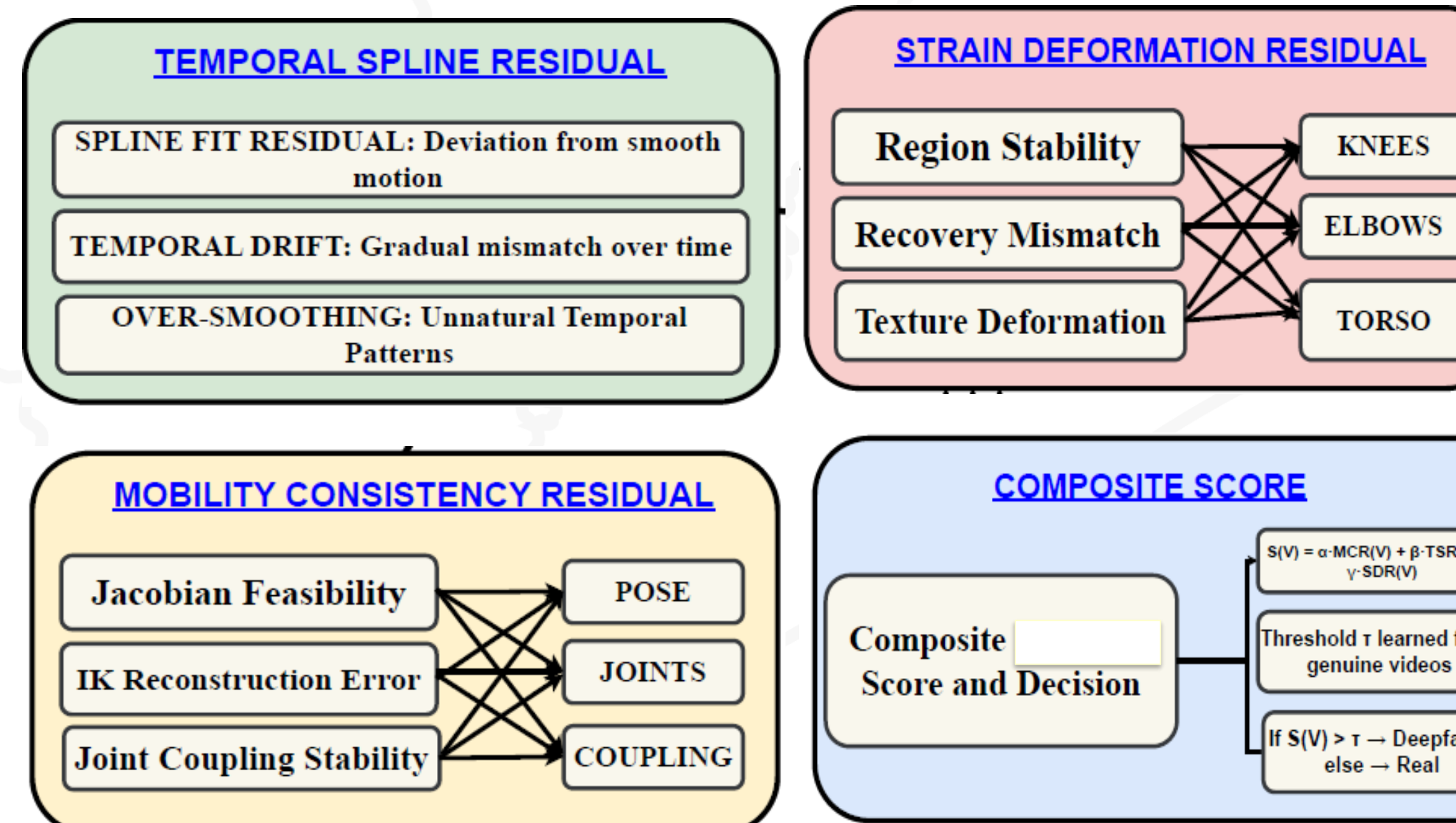
Truth: Looks Real, Moves Wrong



Key Scientific Contributions

Component	Contribution
Formulation	Identity defined via biomechanical motion consistency
Measurement	Motion deviations quantified using TSR and SDR residuals
Principle	Decoupling of appearance and motion patterns
Framework	Identity verified by comparing motion with Human Digital Twin (HDT)
Decision	Detection via residual-based scoring: $S(V) = \alpha \text{TSR} + \gamma \text{SDR}$

A Reverse HDT Framework for Identity Verification



- We analyze biomechanics using temporal coherence and deformation behavior to identify inconsistencies in human movement.
- Temporal spline residual captures motion smoothness over time, while strain deformation residual detects instability in high-motion regions such as joints and facial areas.

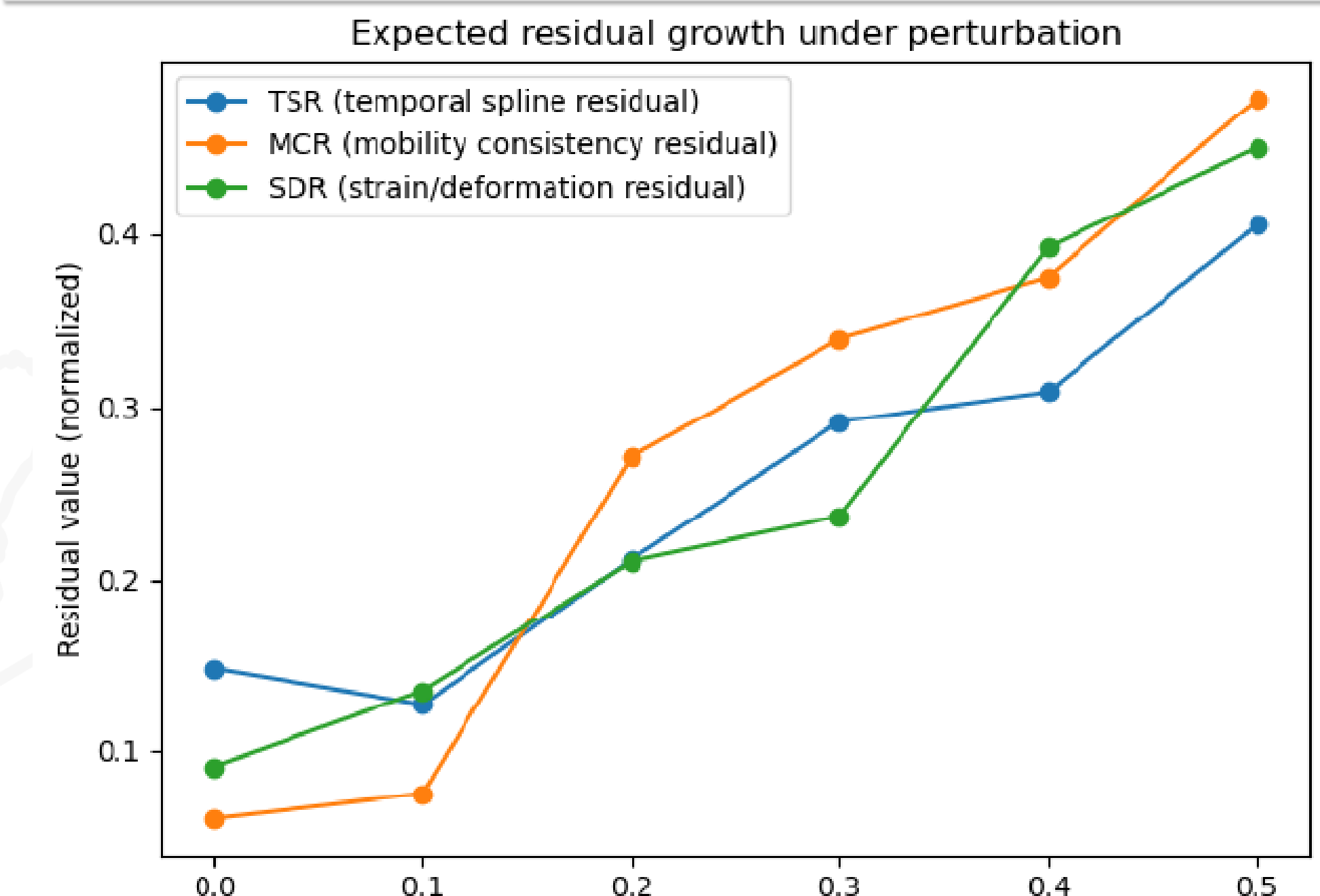
Residual – Based Formulation

$$S(V) = \alpha \cdot \text{TSR}(V) + \gamma \cdot \text{SDR}(V)$$

$$\text{TSR} = \|y(t) - \hat{y}(t)\|^2$$

The detection score $S(V)$ combines temporal and deformation-based residuals to identify behavioral inconsistencies in a video. TSR captures deviation between observed behavior and smooth predicted movement, revealing temporal irregularities.

Residual Trends and Detection Performance



Residual values increase with perturbation strength, indicating growing biomechanical inconsistencies in deepfake-like sequences.

References

S. Talapuru and et al., "Bringing life to 3d human avatars: Integrating motion and texture in real-time," in Proceedings of the 30th International Conference on 3D Web Technology, 2025.

Acknowledgment

