Evaluation of Shape Description Metrics applied to Human Silhouette Tracking

Olivier St-Martin Cormier and Frank P. Ferrie
Dept. Electrical and Computer Engineering & McGill Centre for Intelligent Machines

Introduction

We describe a framework to experimentally compare commonly used shape similarity metrics to determine which is better suited for the task of tracking models with high number of degrees of freedom.

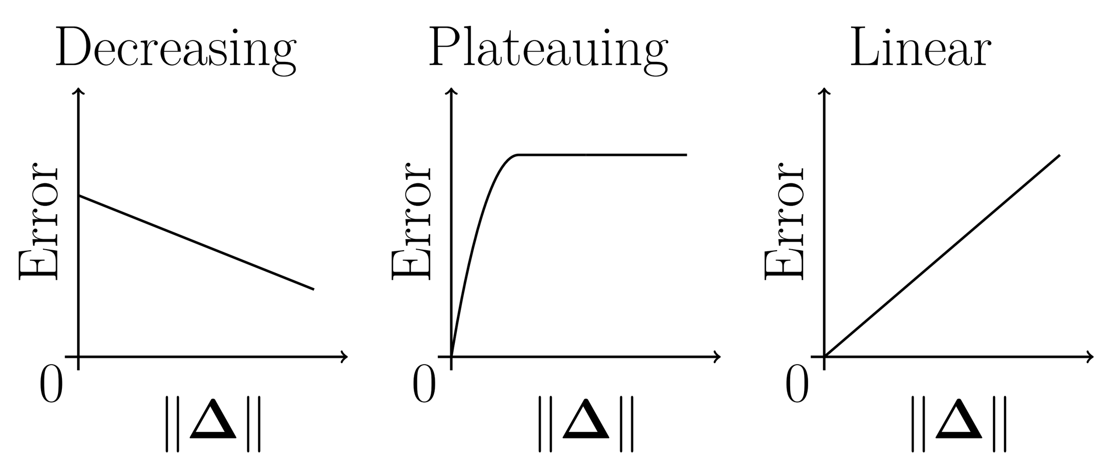
Primary Objectives

- Describe appropriate metrics
- Define a metric evaluation methodology
- Execute experiments to collect data
- Rate the metrics relative to each others

Desirable Metric Characteristics

- Increases monotonically with distance
- Robust to noise and shadows
- Corrolates well with distance in pose space

Expected Metric Behaviors



Evaluated Metrics

Value-based Metrics

- **Hu Moments:** Represents basic properties such as area, centroid, and orientation
- **Pixel Count:** Represents the number of pixels that differ between silhouettes

Edge-based Metrics

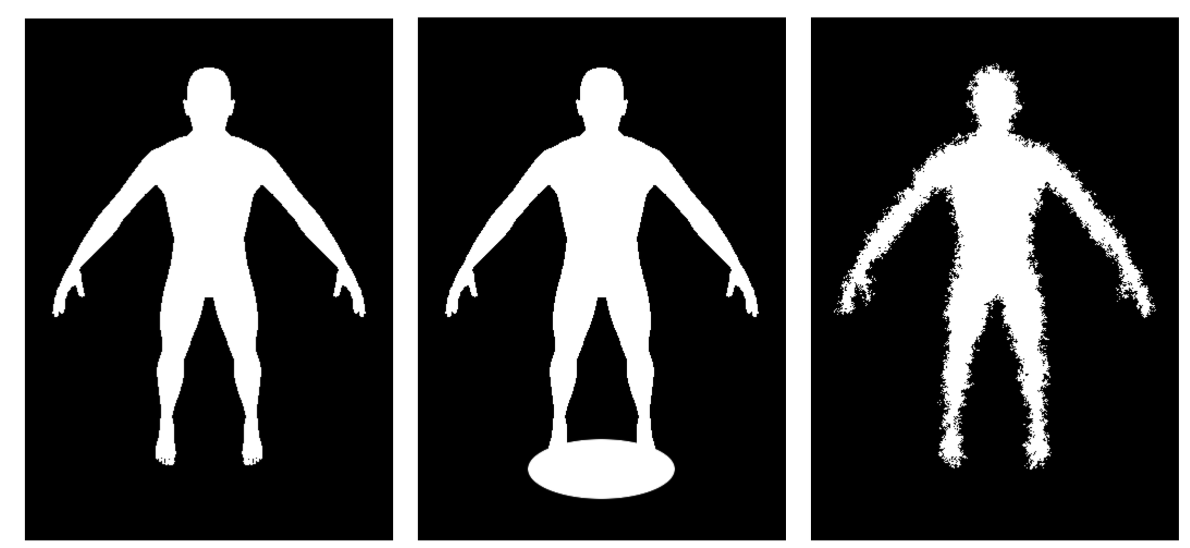
- **Chamfer Distance:** Distance computed with modified Haussdorf distance between the chain codes
- **Turning Angle:** Uses the angle between successive points in the chain code to represents local curvature Distance computed as sum of squared differences between the turning angle representations of each silhouettes
- **Distance Signal:** Uses the relative position of chaincode points to the center of mass of the silhouette Distance computed as sum of absolute differences between the distance signals
- **Shape Contexts:** 2D histograms of relation between each chain code point and other points in the chaincode Distance computed by using a χ^2 test between matched sets

Methodology

Experimental System

- 37 DOF deformable human model
- OpenGL renderer with silhouette shader
- We assess the metrics under 2 different orientations and 3 cases

Test Cases



- We record metric values as the distance in pose space is increased

Conclusion

- Hu moments are more suited to distinguish between classes of shapes
- Turning angles are not robust to noise and shadows
- Distance signals are not robust to shadows
- The chamfer distance has the largest average monotonic region
- The pixel count has the largest correlation coefficient to the distance in pose space

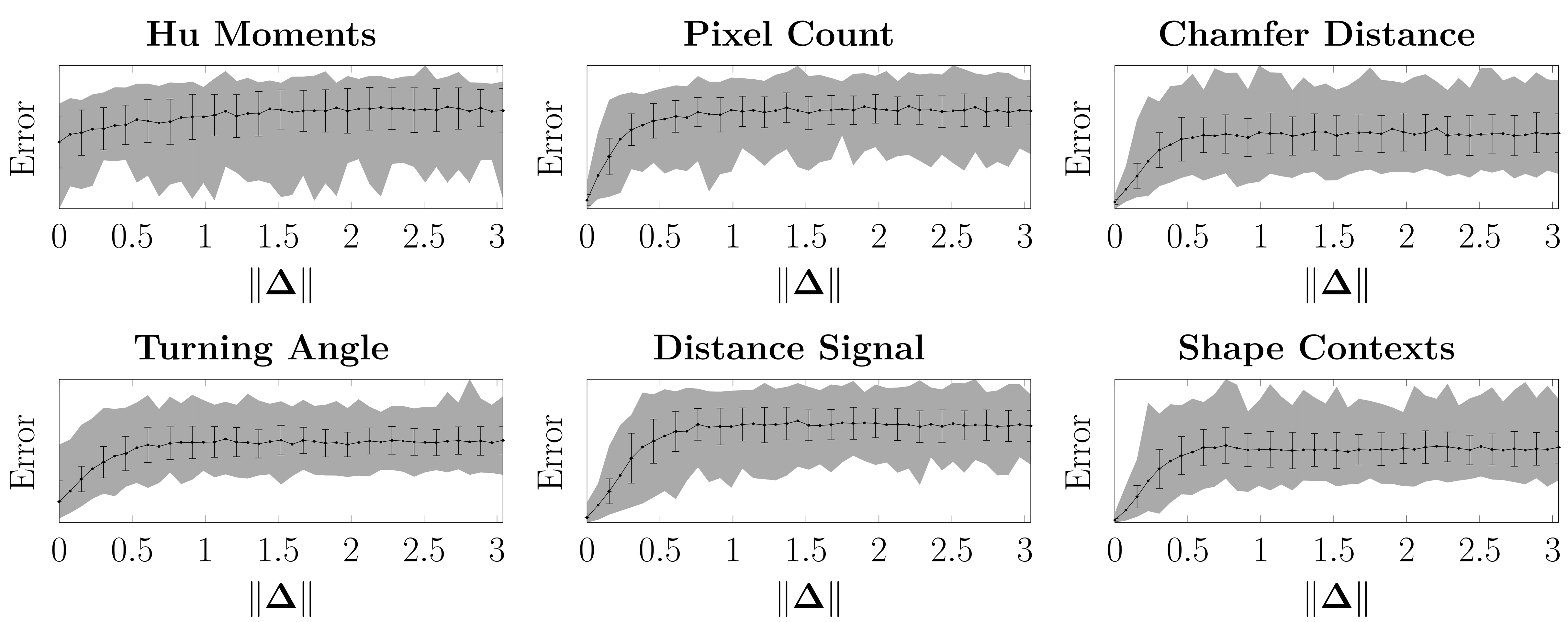
Key Findings

- All metrics studied are functional in ideal conditions
- The pixel count, chamfer distance, and shape contexts are robust to perturbations encountered in human tracking applications
- Because of lower computational cost and higher correlation, the pixel count metric is deemed superior

Contact Information

- http://cim.mcgill.ca/~olivier/CRV2016/
- {olivier,ferrie}@cim.mcgill.ca

Sample Experimental Results



Summarized Results

	Clean				Floor Shadow				Noise				Run Time (μs)	
	Monotonic Region		Correlation Coefficient		Monotonic Region		Correlation Coefficient		Monotonic Region		Correlation Coefficient			
Metric	Front	Side	Front	Side	Front	Side	Front	Side	Front	Side	Front	Side	Average	Std.Dev.
Hu Moments	0.53	0.38	0.168	0.337	0.08	0.30	0.117	0.201	0.08	0.38	0.039	0.338	93.70	19.60
Pixel Count	0.61	0.46	0.730	0.746	0.61	0.46	0.747	0.741	0.61	0.76	0.727	0.748	0.59	0.08
Chamfer Distance	0.61	0.76	0.688	0.718	0.61	0.76	0.686	0.709	0.61	0.76	0.681	0.714	4300.88	1071.97
Turning Angle	0.61	1.37	0.672	0.700	0.00	0.61	0.415	0.666	0.08	0.76	0.099	0.572	37.22	6.28
Distance Signal	0.76	0.76	0.796	0.743	0.00	0.76	0.208	0.658	0.76	0.76	0.797	0.718	0.42	0.06
Shape Contexts														
Greedy Matching	0.61	0.46	0.714	0.684	0.61	0.46	0.691	0.659	0.61	0.46	0.703	0.681	30.75	6.86
Bipartite Matching	0.61	0.46	0.744	0.692	0.61	0.46	0.728	0.664	0.61	0.46	0.726	0.677	38.07	6.73