Lions and Tigers and Bears: Investigating cues for expressive creature motion

Meredith McLendon * Ann McNamara† Tim McLaughlin‡ Ravindra Dwivedi§
Department of Visualization - Texas A&M University

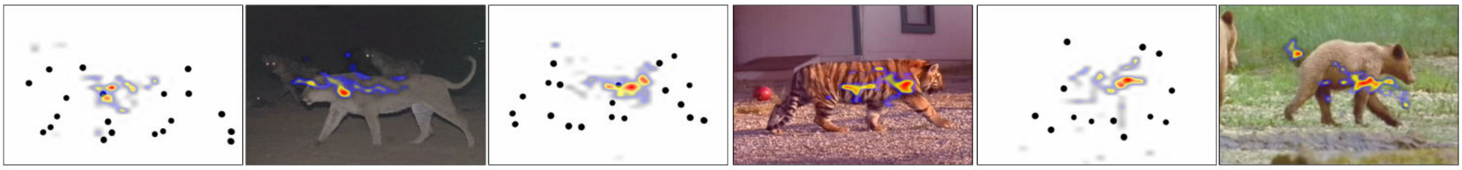

Figure 1: *A single frame, from left to right, eye-tracked data over full representation and over the Point-Light Display.*

1 Perception of Point Light Displays

A digital creature's performance can be thought of as a combination of specifically defined motion and form; a combination that allows the viewer to comprehend the creature's action and intent. Computer graphics offers a variety of methods for defining motion including key-frame animation, data-driven action, rule-based and physically-based motion. However, all of these methods can be complex and time-consuming to implement. Essentially, most computer animation methods force the animator to think about motion at a low-level of abstraction. To create animation tools that simplify the process of creating expressive motion, we need to allow animators to work at a high-level of abstraction. We need determine the minimal elements of form and motion that visually communicate a maximal amount of information about an actor's identity or intentions. By attaching small reflective objects to joint pivot locations and recording at high contrast [Johansson 1973] developed a method for isolating motion from form as a collection of particles, now commonly known as a Point-Light Display (PLD). Manipulating this minimized visual information can even affect the perceived gender of PLD walkers. Cutting [1978] found that exaggerating the movement of points representing the hips and shoulders can bias gender recognition. The goal of our study was to investigate whether viewers use similar visual information to recognize expressive characteristics in animal motion PLDs as when viewing full representations and discover how it might be possible to use that visual information to influence the viewer's perception.

2 Experiment & Results

Our experimental design was inspired by the work of Mather and West [1993] using animal motion captured in side view, where the kinematic motion is most apparent. We used 30 video sequences, creating a corresponding PLD representation for each animal's motion. Three separate experiments were conducted, each focusing on recognition of a single trait pair: (1) Heavy or Light; (2) Predator or Prey; (3) Young or Old. Three groups of five participants, randomly assigned, were eyetracked while viewing the videos. After viewing each video, participants were asked to make a judgment on a five-point Likert Scale with either word from the trait pair at the extremes of the scale. After viewing all of the PLD videos the same process was repeated for the full video presentations. The participant's responses can be analyzed in terms of both *correctness* and *agreement*. We compared each animal species' average mass against the average mass of a human to determine the correctness for Experiment 1. The PLD representation was correctly identified as heavy or light 30% of the time, and identification of the full representation achieved 52% correctness. For Experiment 2, we judged correct identification as a predator or prey by determining an animal's biological classification as a member of order Carnivora. This resulted in a 50% correctness rate for the PLD representation and 87% for the full representation. Judging correctness for Experiment 3 was nearly impossible for all 30 videos since there was no way to know each animal's age. Three videos contained obviously juvenile animals, but only one of the three was correctly identified as young.

Results from a Wilcoxon rank-sum test show that for 93% of the animals in Experiment 1, 87% in Experiment 2, and 70% in Experiment 3, there is a statistically strong probability that the results from both presentation types have the same distribution of responses. This means that similar information is conveyed using both PLD an full presentation. To investigate where this information is concentrated within the videos, we examined fixations recorded using an eye tracker. While nearly all of the fixations in the PLDs occured within our defined regions of interest, as would be expected due to the extreme contrast in the PLDs, 15 of the full videos had 25% of their fixations occur in the regions of interest with five videos matching 70% or better. Further analysis of the eye gaze data for animals that produced both correctness and agreement revealed that participants focused on similar regions in both the full videos and the PLD representations, with much of the gaze locations occuring near the head and shoulder regions. In future research, we plan to isolate these regions and analyze their ballistic information.

Results from these experiments are guiding the development of a framework for an animation system in which the fundamental descriptors of motion are expressive identity cues: minimally defined structures of form and motion which reliably communicate character traits.

*mgmiller@viz.tamu.edu
†ann@viz.tamu.edu
‡timm@viz.tamu.edu
§ravindra@viz.tamu.edu

References

CUTTING, J., PROFFITT, D., AND KOZLOWSKI, L. 1978. A biomechanical invariant for gait perception. *J. Experimental Psychology: Human Perception and Performance 4*, 3, 357–372.

JOHANSSON, G. 1973. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics 14*, 2, 201–211.

MATHER, G., AND WEST, S. 1993. Recognition of animal locomotion from dynamic point-light displays. *Perception 22*, 7, 759–766.

SIGGRAPH 2010, Los Angeles, California, July 25 – 29, 2010.
ISBN 978-1-4503-0210-4/10/0007