Perception of motion *

Vasumathi Raman
Undergraduate Researcher

Jessica K. Hodgins
Faculty Mentor

Final Report
Graphics Lab
Carnegie Mellon University
Sep 1, 2005

* A series of ongoing research projects that Vasumathi Raman contributed to during an internship under Jessica Hodgins at the Graphics Lab, Carnegie Mellon University, Pittsburgh, PA. Student participation supported by the CRA Committee on the Status of Women in Computing Research (CRA-W), via the Distributed Mentor Project (DMP), Summer 2005.

Abstract

The human perceptual system makes a number of assumptions and inductive inferences to resolve depictions of motion. Different features of the image depicted are assigned different weight and considered at different levels while recovering reality from representation. In producing animations of such motion, it is thus to our advantage to consider the assumptions made and the weighting of different aspects of the motion depicted. Our projects concern the perception of human motion, with the goal of uncovering guidelines for producing more compelling animations at minimum cost. Our first project questions the basis of recognition of human motion, and aims to uncover guidelines for what classifies a motion as human and aids its recovery from abstract representation (as dot patterns). Our second project studies the effects of scaling motion with size. We would like to show that scaling motion correctly with size preserves the believability of the animation.

The results we hope to achieve from this research will aid the development of animation strategies that exploit the workings of the human perceptual system to reduce production costs.

Keywords: Johansson point-light displays, occlusion, scaling laws,

1 PROJECT 1 - UNUSUAL HUMAN MOTION

1.1 Introduction

Our first project involves recovery of the human form from dot pattern movies of human motion. It is well known that the human form can be easily recovered from dot-pattern images of a running person. What we are interested in is when this ability breaks down. What determines whether a given dot-pattern is recognizable as human? We believe that this ability has to do with the physics of the motion depicted - non-intuitive motions which appear to violate the laws of physics cause confusion and are less likely to result in successful detection of the human form. We hypothesize that it should be pretty hard to tell exactly what is being depicted when the motion is unfamiliar and unrealistic.

1.2 Background

A number of studies using Johansson point-light displays have shown that subjects can detect walking and running motions, identify the gender of the runner and even recognize the gait of friends simply by observing dot-pattern renditions of these motions.

Proffitt & Bertenthal [1] say that motion information is a minimal stimulus condition for the perception of form, and detail two processing models for displays of moving jointed objects:

1. A top-down approach, which says that the subject takes the human form as given and seeks to match the presented figure to the known model. This model lacks generality, as it is specific to displays of human walkers.
2. A bottom-up approach, which works on a "fixed-axis" assumption and seeks to discover connectivity in the presented pattern by employing a set of assumptions about the motion of rigidly related points rotating in different fields of depth. This model seeks to recover rigid relations by testing whether each pairing of points is

rigidly translating or rotating about an axis fixed in direction. Whenever a pair of points meets this assumption, they are interpreted as rigidly connected. After deriving a set of pair wise connections, this model then proceeds to connect pairs having one point in common.

The top-down approach would suggest that there is something special about human motion that causes it to be so easily identifiable.

1.3 Method

In order to study the role of top-down analysis in recognizing human motion, it would be meaningful to compare results using otherwise identical stimuli that did not represent the human form. Such stimuli would have to incorporate identical motion trajectories with different rigidity constraints. The approach we took was to interchange the motion trajectories of the joints, effectively scrambling the motion while keeping the fundamental nature of the stimulus constant.

There are a few differences between ours models and traditional ones such as those used by Proffitt & Bertenthal. First, we use more points than they did. Second, we use three-dimensional spheres rather than point lights, the implications of which we will discuss later. The dot pattern motions created were based on motion capture of the human running motion, with spheres added at the joints using Maya. To scramble the motion, we obtained the initial locations for each sphere and swapped them at random. We also added some noise to alter the initial positions of the spheres from the human form

We are also trying to study the effects of occlusion on subjects' abilities to identify the motions as human. A previous study conducted by Proffitt, Bertenthal & Roberts [2] used movies of humans with digitized point-lights at the joints, which showed that subjects are more likely to correctly identify a motion as human when all the correct occlusions are present. Occlusion information served two purposes – to provide depth cues and resolve multistability of the displays, and to provide information about the presence and location of occluding objects. However, we use three-dimensional spheres instead of point lights (Fig. 1.3.1), so the question of multistability lessens in importance due to the perceivable occlusion of the spheres themselves. Point light projections were very two-dimensional, with no explicit depth cues available to the observer - when two point lights cross, we cannot tell which is in front. With our model, on the other hand, there are very explicit depth cues available. The only thing missing is the occluding object information.

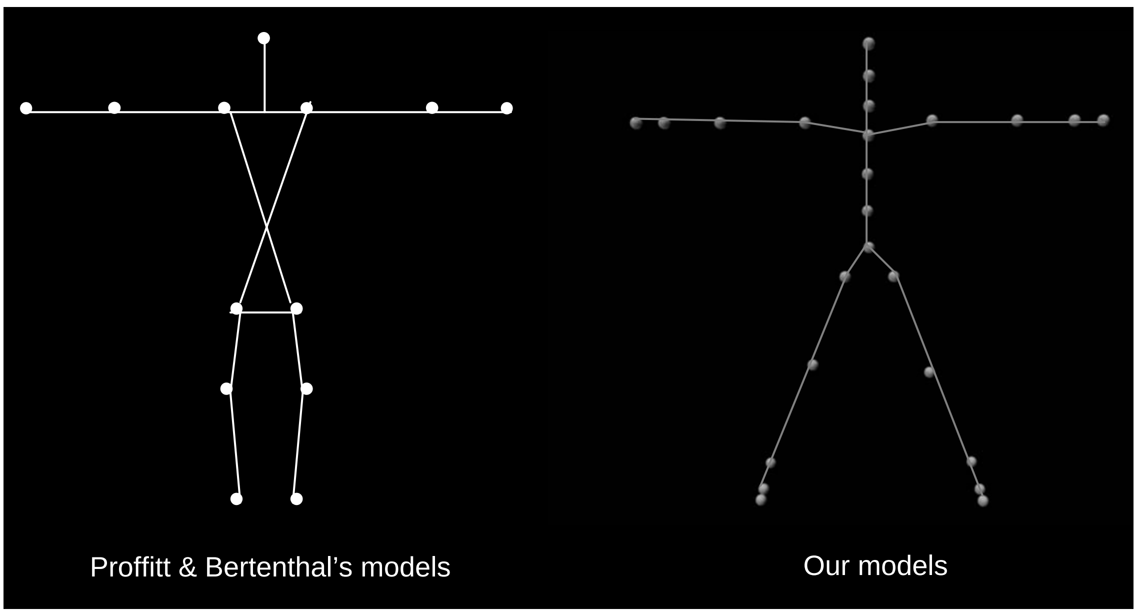

Fig 1.3.1 *Our models vs. traditional models*

To overcome this deficiency, we tried to create camera angle independent occlusion effects using a human shell and back-facing with reversed normals (Fig. 1.3.2). Back-facing is a technique used to reduce rendering time by only rendering those surfaces that face the camera. Reversing the surface normals before back-facing allows us to render only those surfaces facing away from the camera, thus ensuring that the spheres within are always seen.

This was a challenging task, and we had to greatly modify the shell in order to achieve the desired effects. In spite of this effort, the resulting motions were even less recognisable than before and the final movies did not feature occlusion.

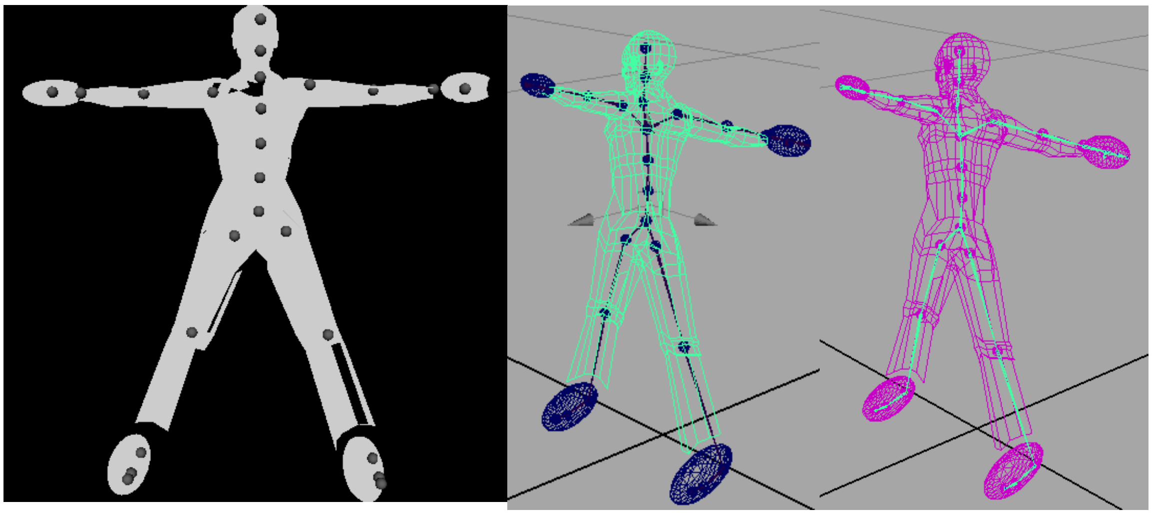

Fig 1.3.2 *Occlusion using a back-faced human shell*

In addition to the scrambled running, we developed another set of motions to test our hypothesis of physically unusual or impossible motions being less easily recognised. These included dot-pattern renditions of two people forming a human wheelbarrow, a couple holding hands and spinning around and a pair of contact improvisation dancers. The common feature of these motions was that separating out the motion for a single member of the pair performing the motion (e.g. the barrow without its driver) produced a motion that would be physically impossible on its own.

We sent these movies to our collaborators at the University of Virginia for their opinions on the types of experiments that could be conducted, but did not hear back from them before the summer's end.

2 PROJECT 2 – SCALING LAWS AND ANIMATION OF MOTION

2.1 Introduction

Another project we worked on uses scaling laws in animating motion. Intuitively, the size of a creature affects how we expect it to move. A smaller creature is expected to scamper and a larger one to lumber. It is believed that motion, if scaled with size, "breaks", i.e. ceases to be believable. We want to show that motion, if scaled correctly, can be preserved. We predict that if we compare subjects' responses to motions scaled with deference to scaling laws and without, we will find that the incorrectly scaled motions break down while the correctly scaled ones remain cohesive.

2.2 Background

A number of psychological studies have shown that scaling motion proportional to a change in size is not sufficient to preserve its credibility. Our experiments make use of scaling laws used by Raibert & Hodgins [4] to scale motion in response to a scaling in size. Motion is preserved as long as the perceived gravity remains true to its real value of 9.8 ms^{-2}. Thus, scaling size (in m) by a factor of x requires the time taken to complete the action to be scaled by a factor of $(1/\sqrt{x})$ to conserve gravity. Any other scaling of the motion will cause the perceived gravity to waver from its real value and the motion will cease to be believable.

2.3 Method

2.3. A Movies

Our experiments make use of small, normal (~human) and large sized imaginary creatures called "blobs" (Fig. 2.3.1), with differently scaled motion. The motion here is a non-human sliding across a checkerboard level plane. Motion scaling was achieved by modifying the frame rate. If we scale the blob by x (in m) and the time taken to complete the action by t (in s), gravity (in m/s2) would be scaled by x/t^2. In order to conserve gravity, $x/t^2 = 1$, therefore t is scaled by $(1/\sqrt{x})$ for correct motion scaling. Thus, in the "natural" condition, the smallest blob is the fastest moving and the largest one the slowest.

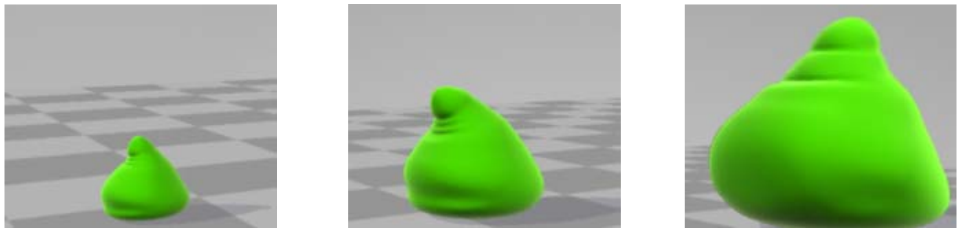

Fig 2.3.1 *Small, normal and large blobs*

We also prepared movies in which the motions were interchanged, i.e. the large blob moved fast, the small blob moved slowly and so on. We thus produced a total of nine movies, with one correct scaling and two incorrect scaling for each of three blob sizes.

2.3. B Survey

We designed an electronic survey to gauge subjects' response to the variously scaled motions. At the beginning of the survey, subjects were shown a still image (Fig. 2.3.2) of the three blobs side by side, with a tree and a deck chair in the background to serve as a size reference, since it was decided that the checkerboard floor alone did not provide sufficient cues to this end. The survey displays the nine movies in random order to minimise errors arising from prejudiced judgements about the later movies based on those seen first. A random ordering of the 3 sizes is coupled with a random ordering of the 3 speeds, producing a total of 36 combinations (possible orders).

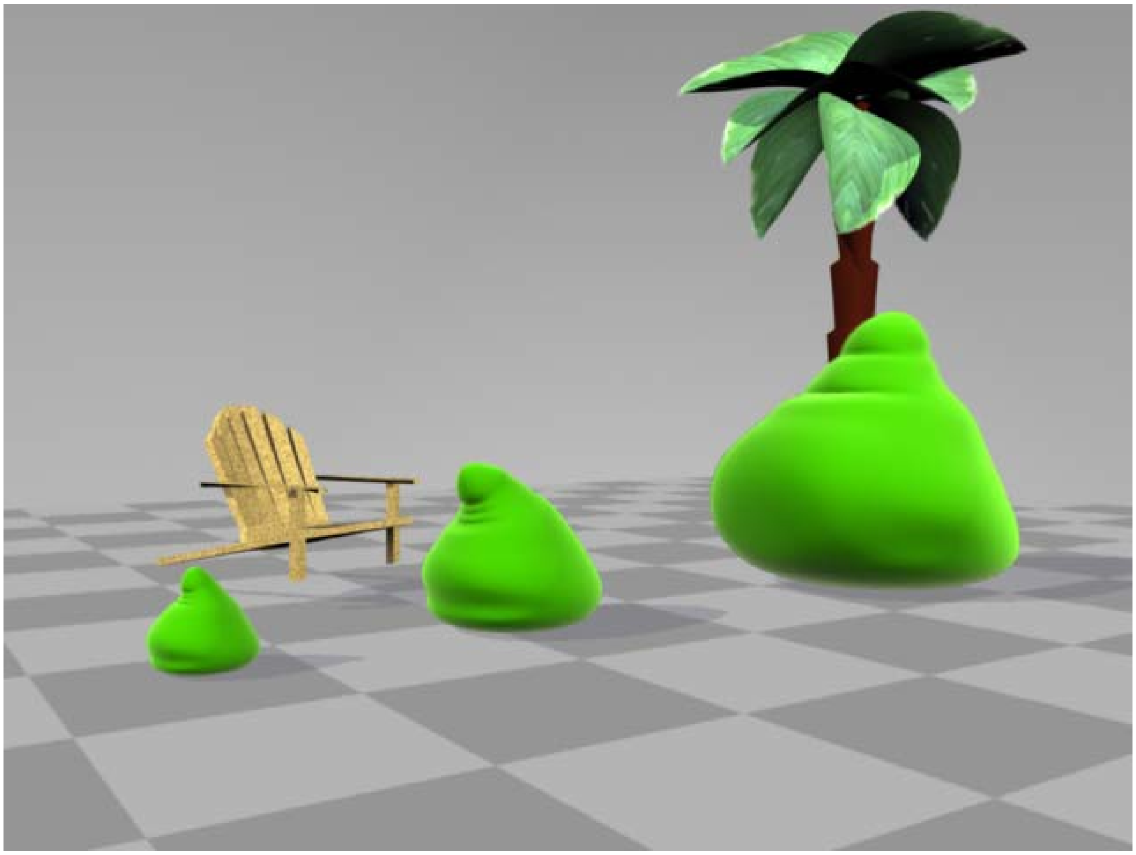

Fig 2.3.2 *Still shown at beginning of survey, with background objects for size comparison*

Participants are asked to fill out a short survey after watching each movie. Questions asked of participants include:

- *Is this glob smaller or bigger than the average person?*
- *Does this glob move slower or faster than a person normally walks?*
- *How unnaturally or naturally does this glob move?*
- *How humanlike does this glob seem?*
- *How smart does this glob seem?*

Answers are obtained on a seven degree Likert scale (Fig. 2.3.3).

About This Glob							
	Smaller						Larger
Is this glob smaller or bigger than the average person?	○	○	○	○	○	○	○
	Slower						Faster
Does this glob move slower or faster than a person normally walks?	○	○	○	○	○	○	○
	Unnatural						Natural
How unnaturally or naturally does this glob move?	○	○	○	○	○	○	○
	Not Humanlike						Humanlike
How humanlike does this glob seem?	○	○	○	○	○	○	○
	Not Smart						Smart
How smart does this glob seem?	○	○	○	○	○	○	○

Fig 2.3.2 *Survey questions with Likert scale responses*

This survey is currently underway. We hope to have results soon.

We are also working on a sub-project that applies scaling principles to human motion, specifically dynamic human motions such as back-flips. The reason we used the back-flip is because the long time spent in the air emphasises the effects of incorrect gravity. We plan is to conduct three separate studies

- the blobs without background objects (but with an initial still for size reference)
- the blobs with a tree and a chair in the background for each movie
- the human performing a back-flip

The first of these is underway. We intend to show that

- scaling motion correctly with size preserves its credibility ,
- adding objects for scale increases this effect, and
- the effect is stronger for human motion than for non-humanoid motion (such as that of the blob).

3 Acknowledgements

My participation in this research was made possible by the sponsorship of the CRA Committee on the Status of Women in Computing Research (CRA-W) via the Distributed Mentor Project (DMP) program, and the generosity of my mentor, Jessica K. Hodgins at the Carnegie Mellon University Graphics Lab.

4 References

[1] Proffitt, D.R. & Bertenthal, B.I. (1988). Recovering connectivity from moving point light displays. In W.N. Martin & J.K. Aggarwal (Eds.), Motion understanding: Robot and human vision, Hingham MA: Kluwer.

[2] Proffitt, D.R., Bertenthal, B.I., & Roberts, R.J., Jr. (1984). The role of occlusion in reducing multistability in moving point light displays. Perception & Psychophysics, 4, 315 323.

[3] Bertamini, M. & Proffitt, D.R. (2000). Hierarchical motion organization in random dot configurations. Journal of Experimental Psychology: Human Perception and Performance, 26, 1371-1386.

[4] Raibert, M.H. & Hodgins, J.K. (1991). Animation of dynamic legged locomotion. Proceedings of the 18th annual conference on Computer graphics and interactive techniques, 349 – 358.