Evaluating Human Motion

Jessica Hodgins
Robotics Institute and Computer Science
jkh@cs.cmu.edu
www.cs.cmu.edu/~jkh

How to evaluate?

- Looks good…
- Side-by-side comparisons (~1995)
- Perceptual Studies (~1998)
  - Hodgins, O’Brien and Tumblin
  - Harrison, Rensink, van de Panne

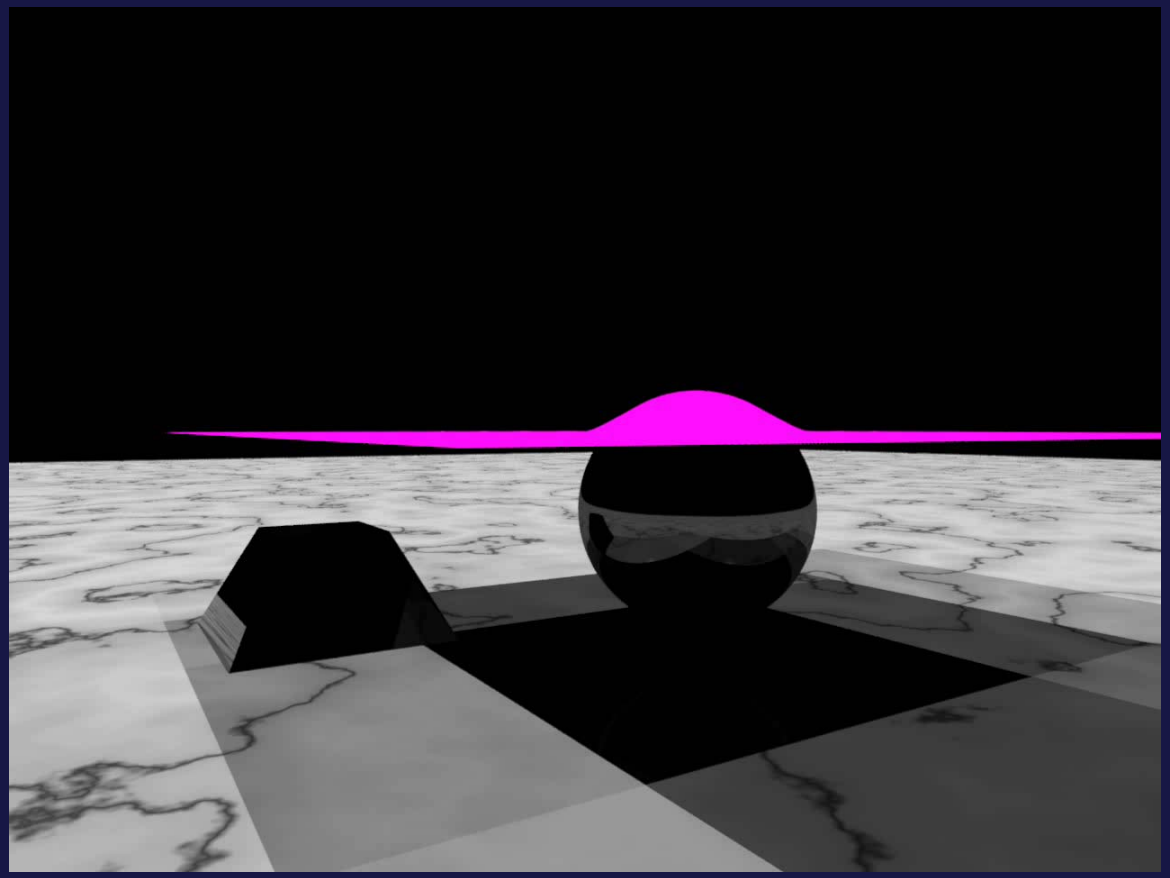
Perceptual Studies

- Noticable difference
  - Both papers
- Survey questions
  - How angry is this person?
- Free-form questions
  - What did you see in this video?
  - What did you like about it?
- Behavioral response
  - Eye tracking
  - fMRI
  - Enactment or Interference

Experimental Setup

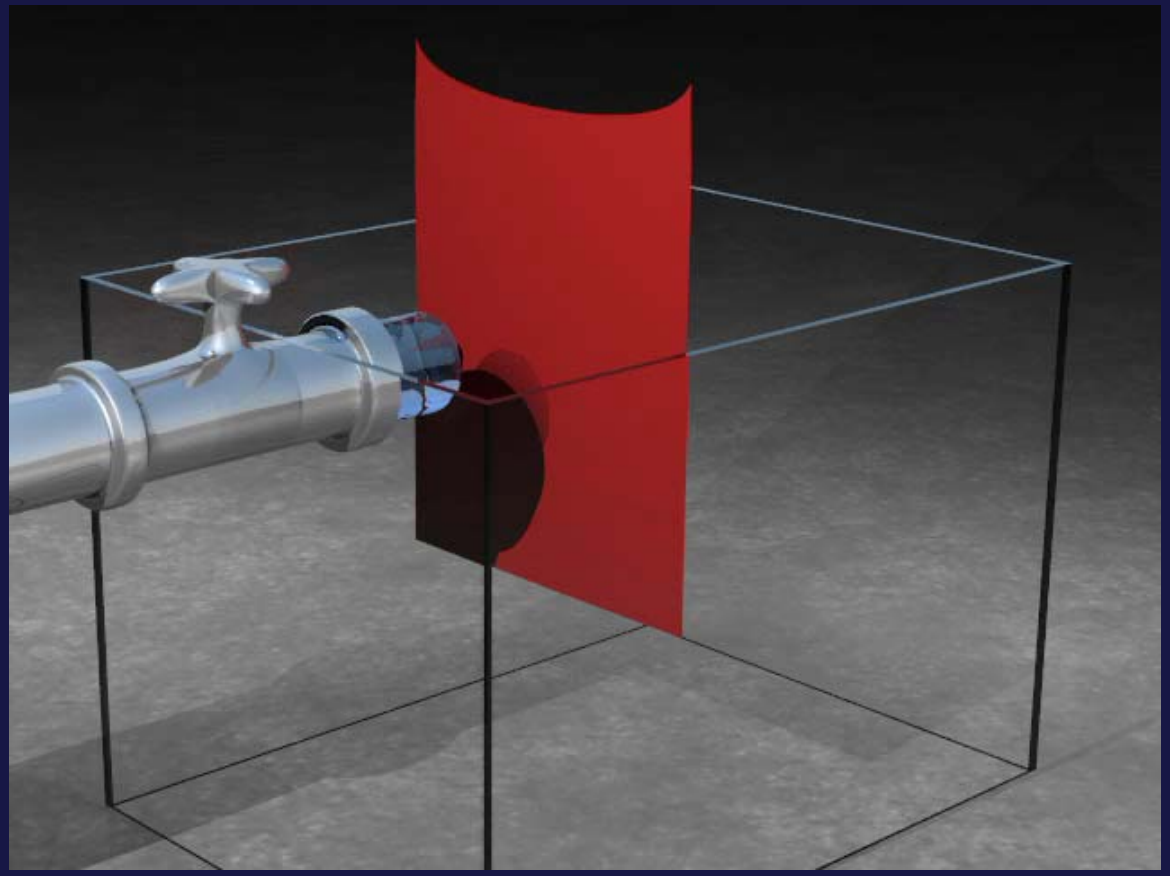
- Rendering style
- Complexity of scene—attention matters
- Distracter Task
- Many others

Looks Good?



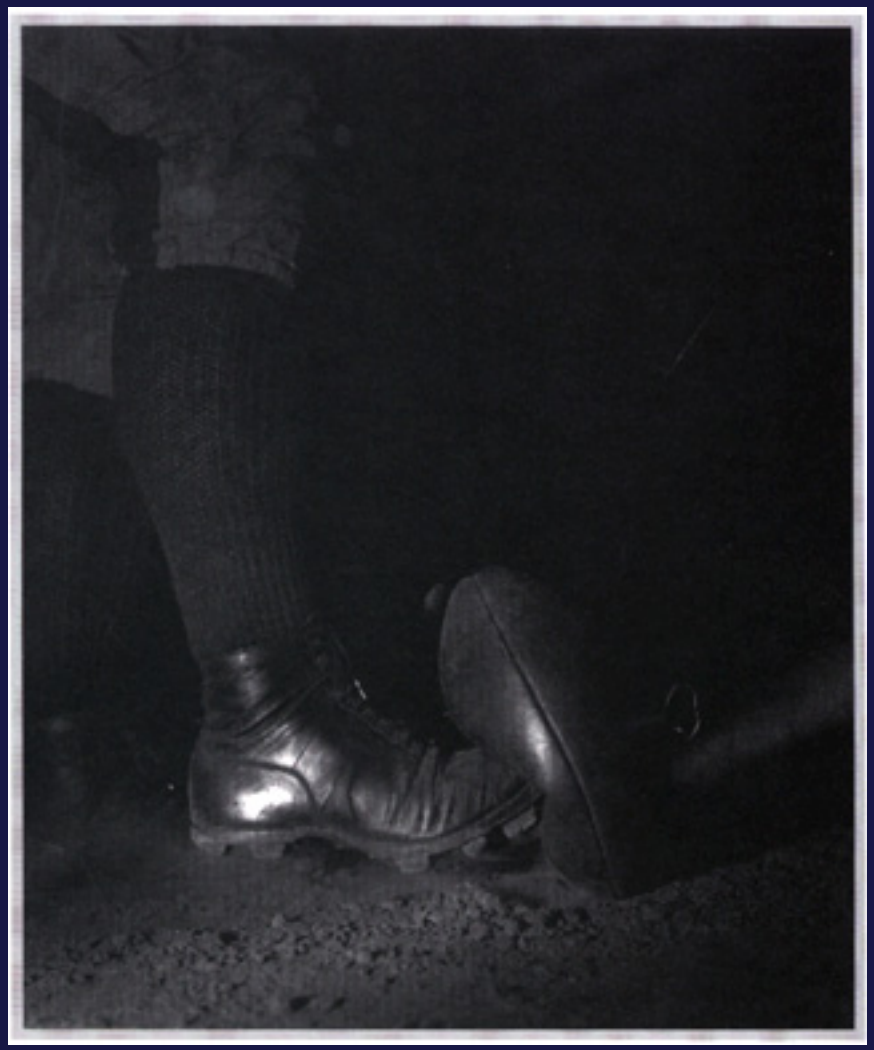
Ron Fedkiw, Robert Bridson, and John Anderson

Looks Good?



Ron Fedkiw, Eran Guendelman, Andrew Selle and Frank Losasso

Looks Good?



Wes Fesler Kicking a Football, 1934

Dr. Harold Edgerton

Looks Good?

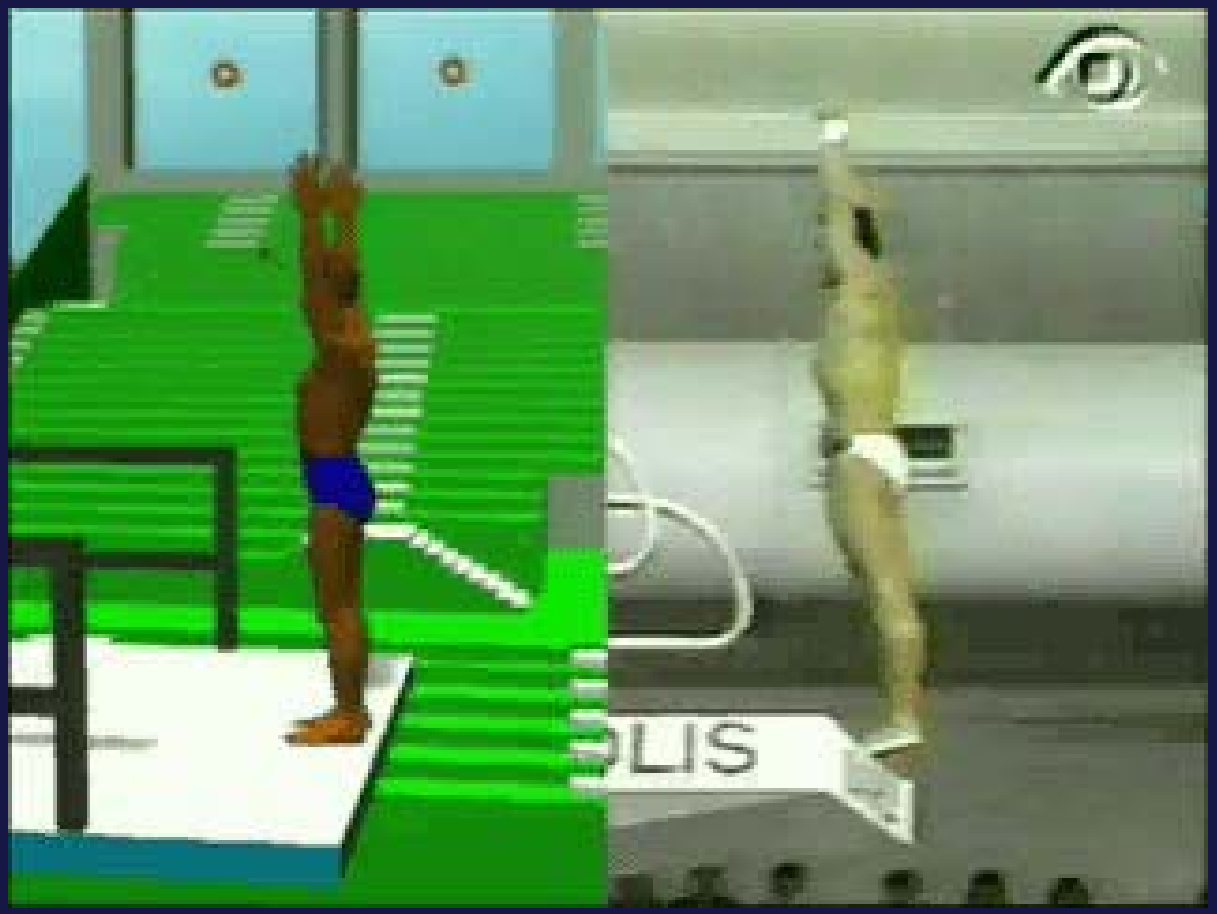
1000 fps

Side-by-side Comparison

Side-by-side Comparison

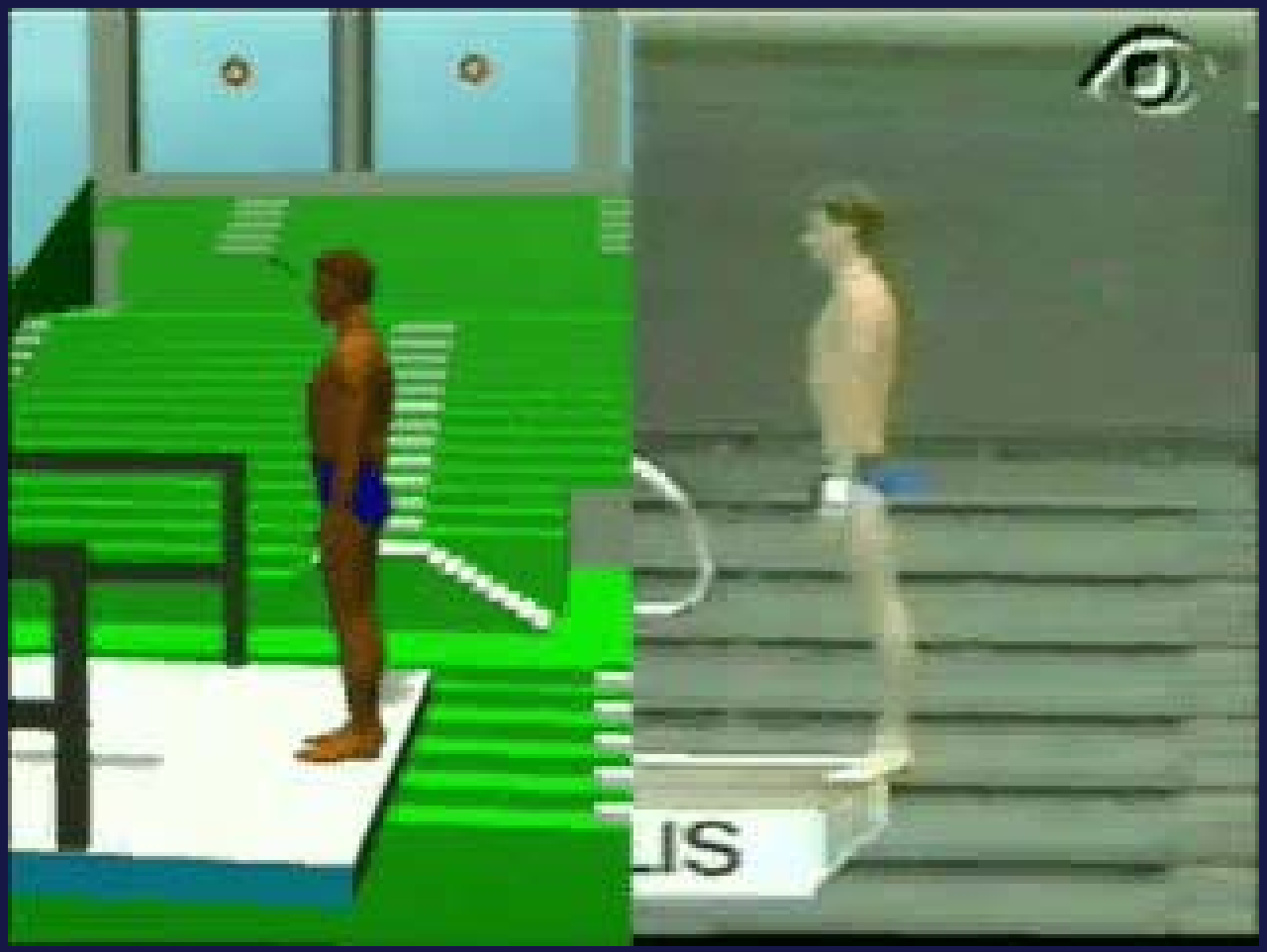
Victor Zordan

Side-by-side Comparison



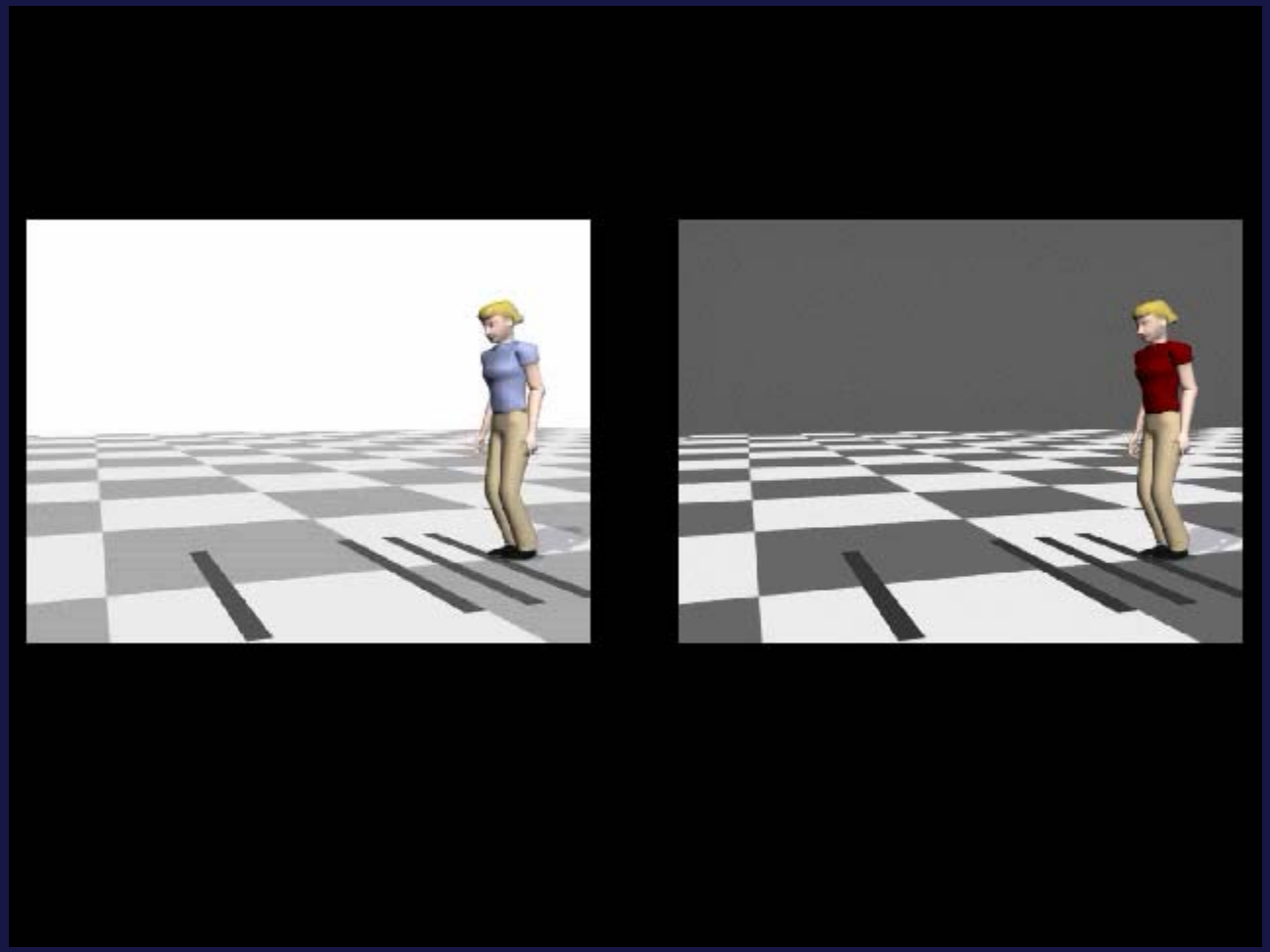
Wayne Wooten

Side-by-side Comparison



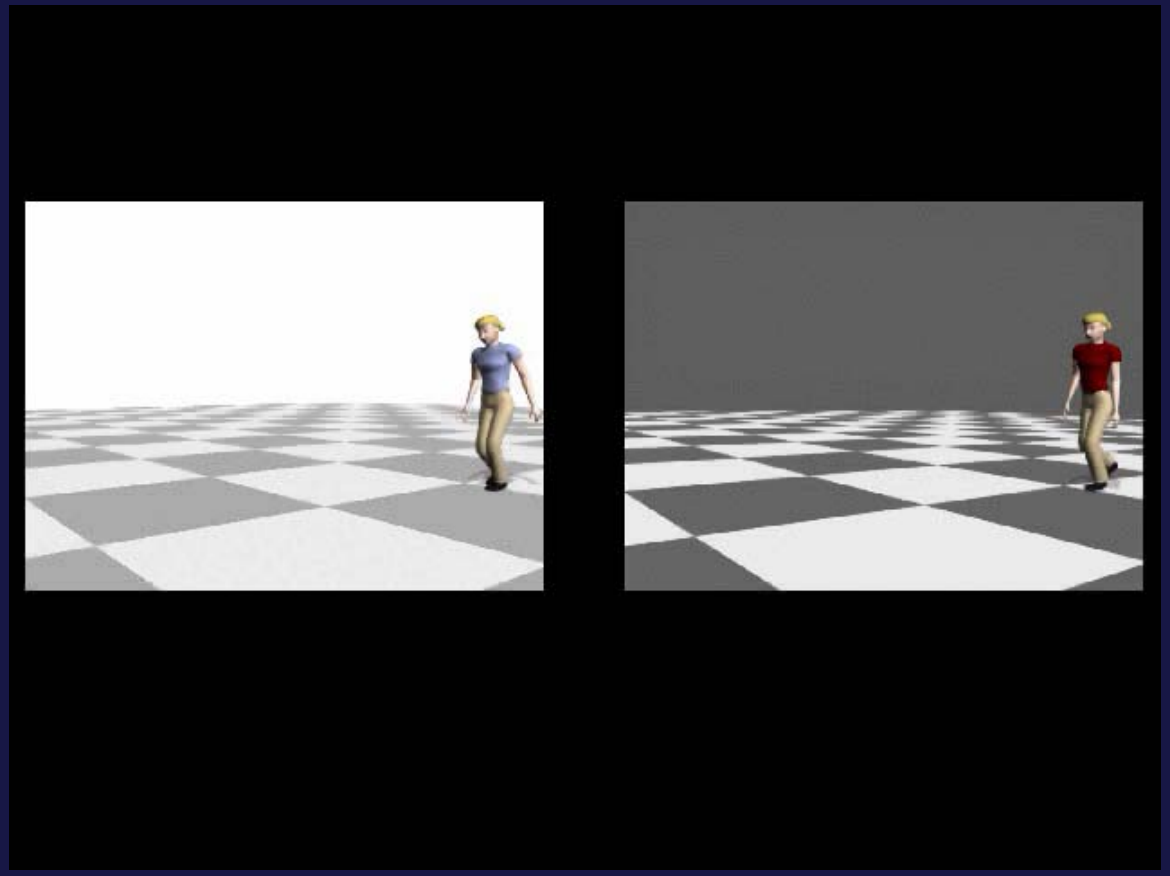
Wayne Wooten

Side-by-side Comparison



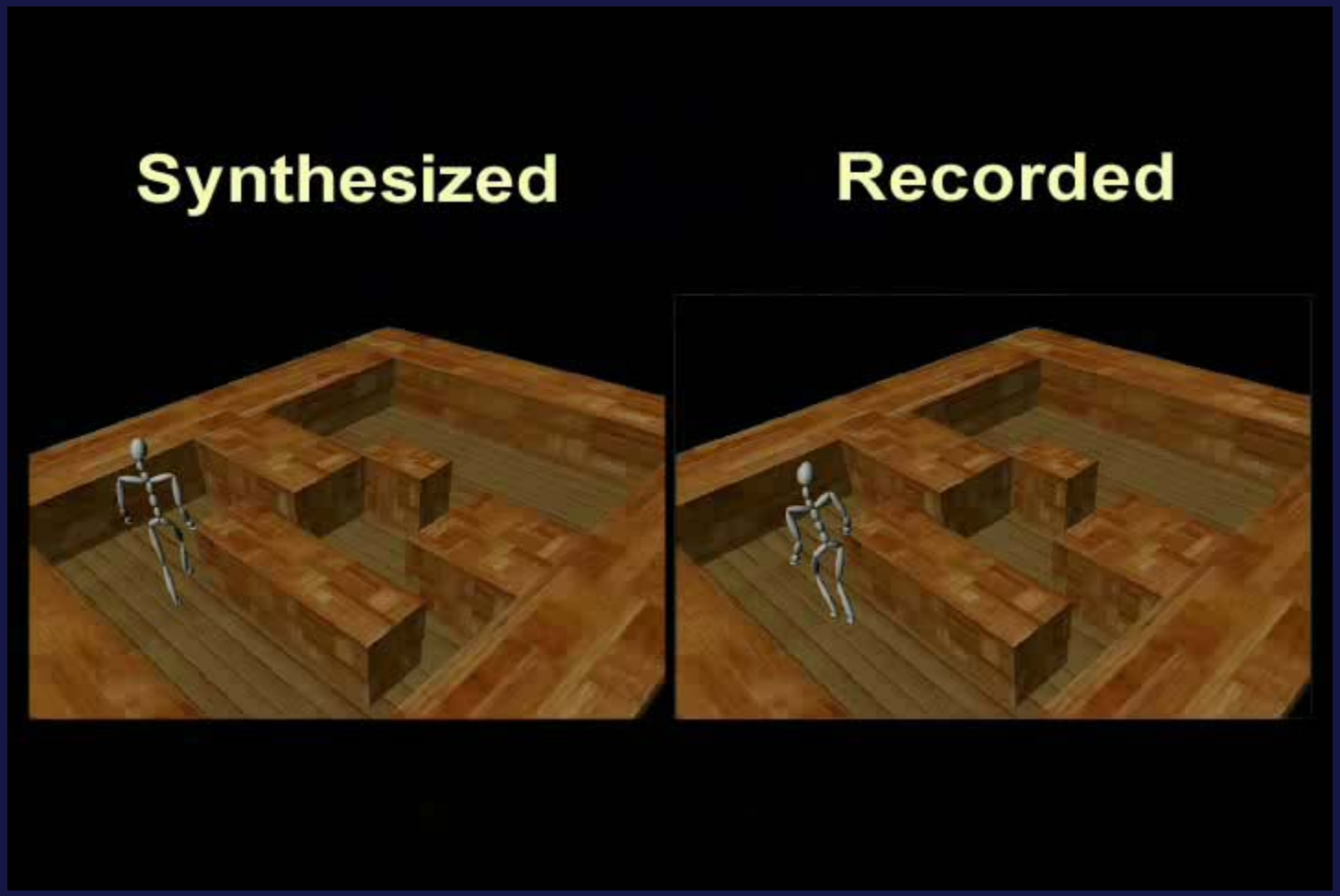
Side-by-side Comparison

Side-by-side Comparison



Side-by-side Comparison

Side-by-side Comparison



Side-by-side Comparison

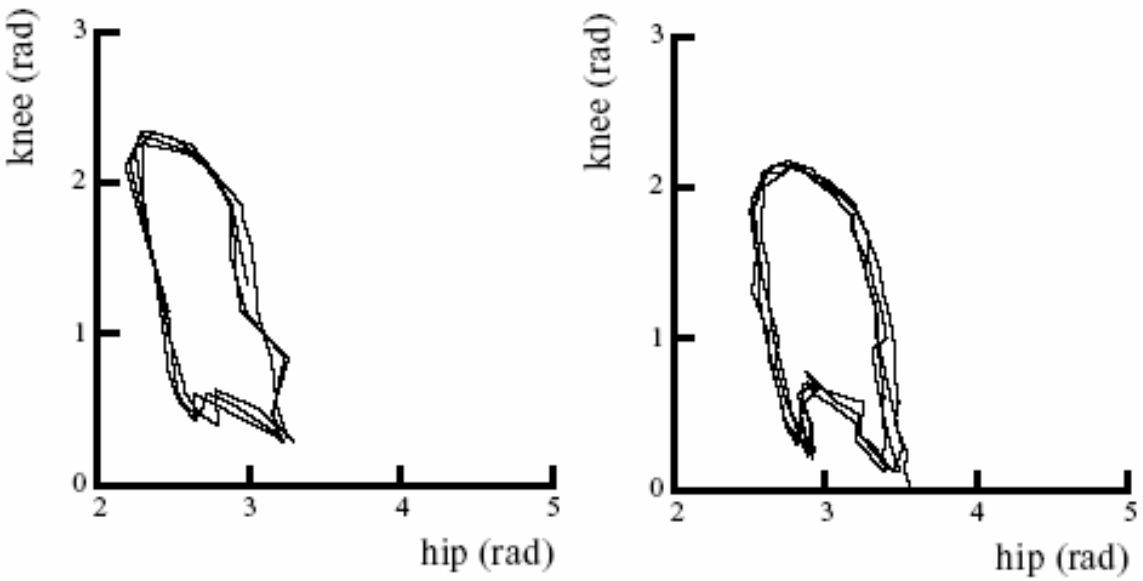


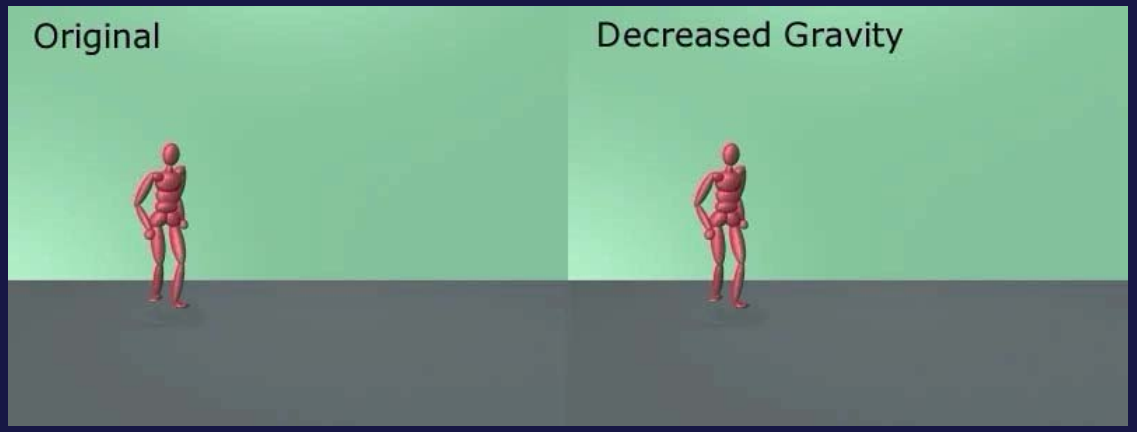
Figure 6: A phase plot of the hip and knee angles seen in the simulated runner (left) and measured in human subjects (right). The simulated motion is qualitatively similar to the measured data.

Or force plate data?

Perceptual Studies

Which motion is wrong?

Can you detect a change?



Reitsma and Pollard, SIGGRAPH 2003

Harrison, Rensink, van de Panne SIGGRAPH 2004

Perceptual Studies

Hodgins, J. K., O'Brien, J. F., Tumblin, J., Perception of Human Motion with Different Geometric Models. IEEE: Transactions on Visualization and Computer Graphics, December 1998, Vol. 4, No. 4, pp. 307-316.

Hypotheses:

- Simple representations → fine distinctions
- Complex, "accurate" representations → fine distinctions
- Equally fine distinctions independent of model

Perceptual Studies

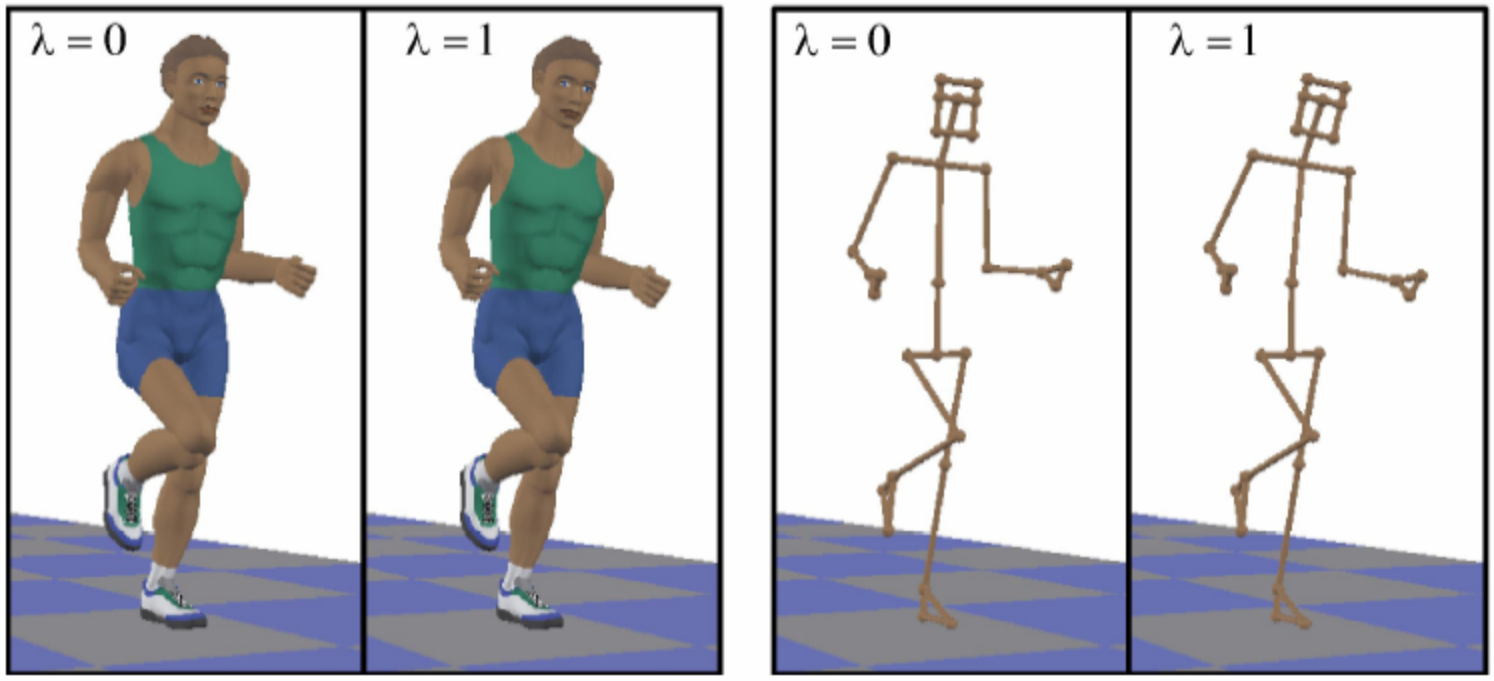


Fig. 1. Images of an animated human runner. (a) Two running motions rendered using a polygonal model. (b) The same pair of motions are rendered with a stick figure model. Modifications to the motion were controlled by a normalized parameter, λ, that varied between $\lambda = 0$ and $\lambda = 1$. These images are from the motion generated for the additive noise test discussed in Section 3.3. The difference in posture created by the additive noise can be seen in the increased angle of the neck and waist in the right image of each pair ($\lambda = 1$).

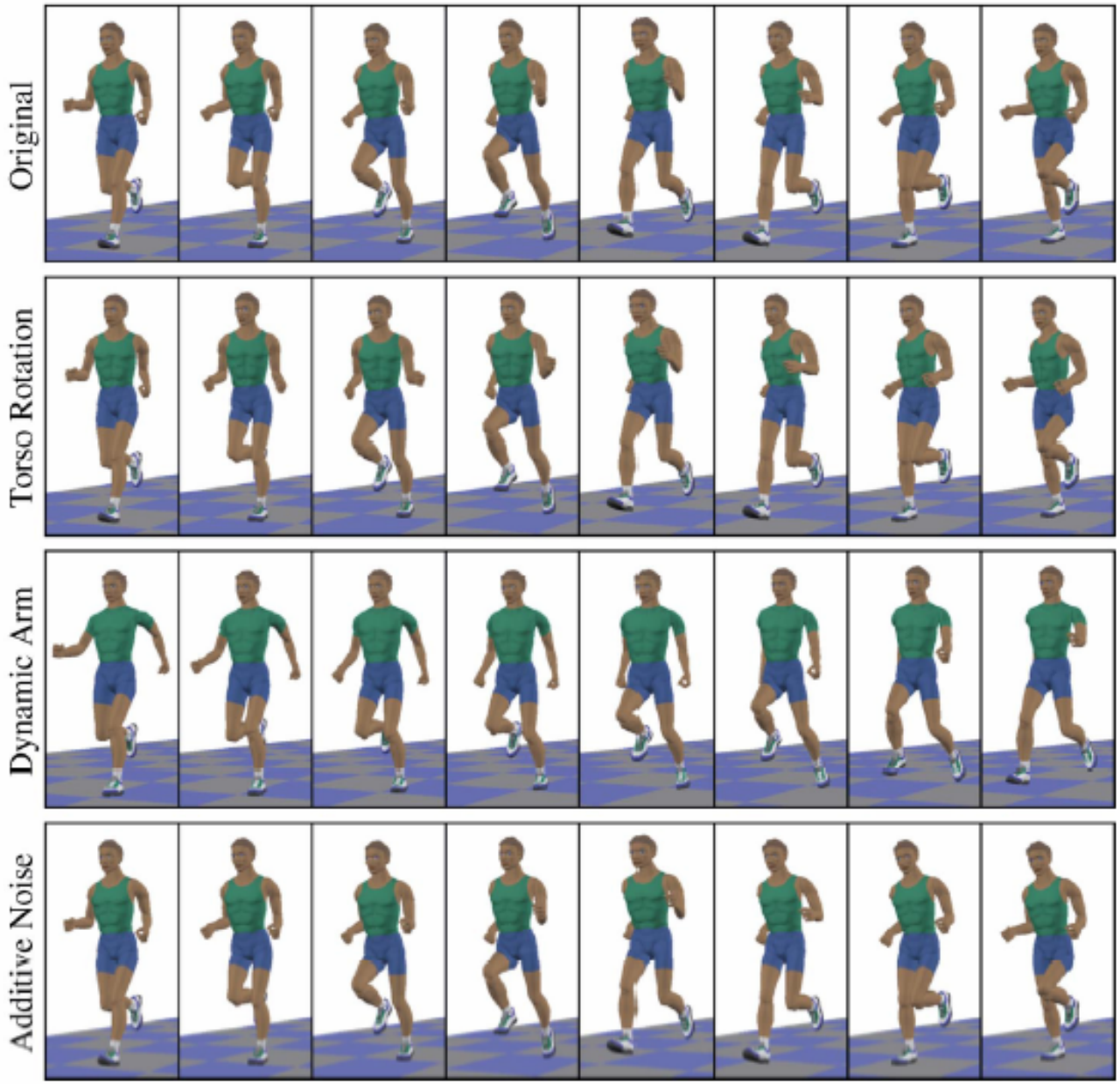


Fig. 3. Examples from the motion sequences rendered with the polygonal model. **First Row:** Original motion sequence, $\lambda = 0$, used in all tests. **Second Row:** Torso rotation motion sequence with 10× magnification of the torso rotation, $\lambda = 1$. **Third Row:** Dynamic arm motion sequence with maximum exaggeration, $\lambda = 1$. **Fourth Row:** Additive noise motion sequence with sinusoidal noise of ±0.15 radians, $\lambda = 1$. Images are spaced at intervals of 0.067 seconds.

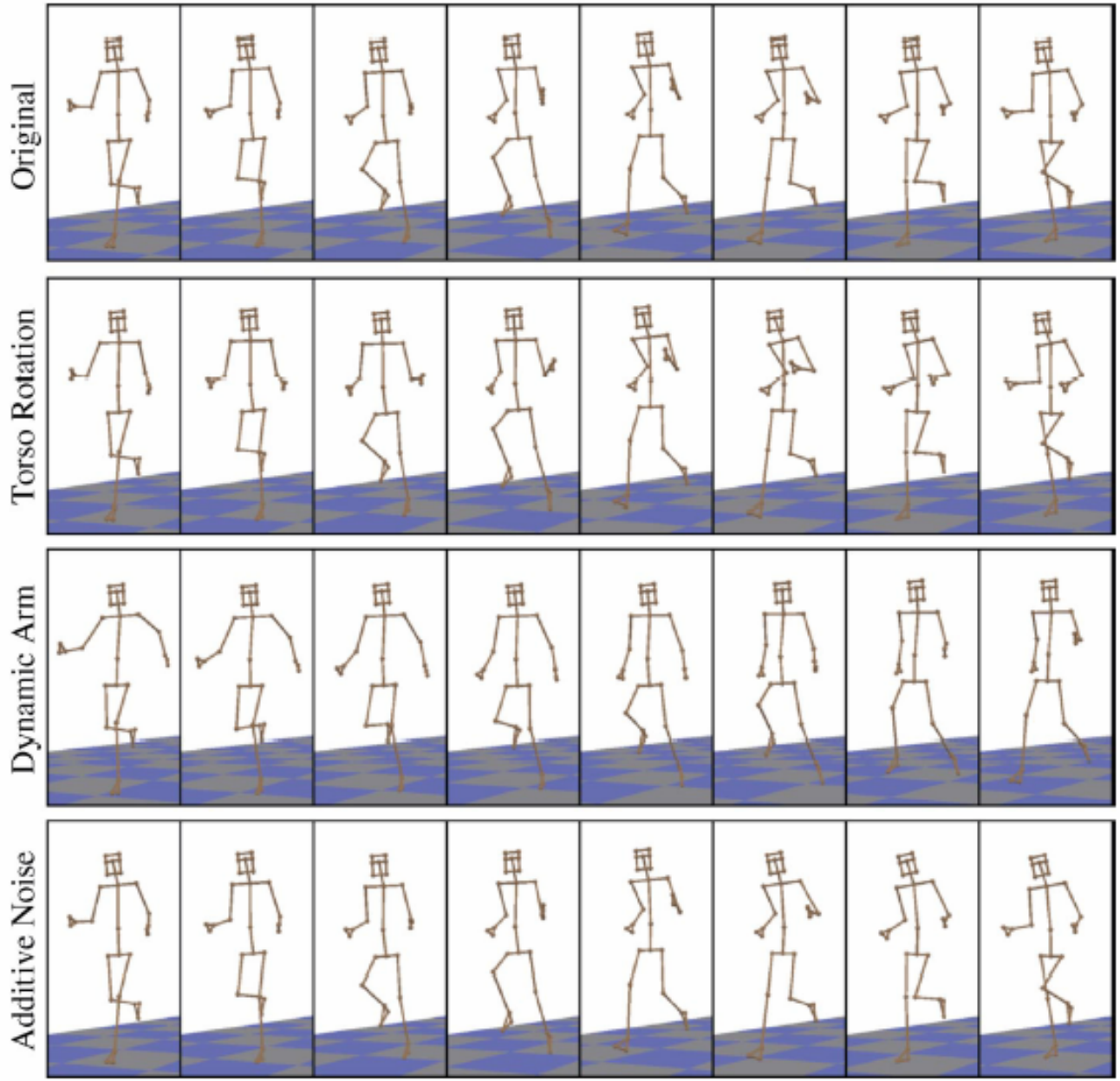


Fig. 4. Examples from the motion sequences rendered with the stick figure model. **First Row:** Original motion sequence, $\lambda = 0$, used in all tests. **Second Row:** Torso rotation motion sequence with 10× magnification of the torso rotation, $\lambda = 1$. **Third Row:** Dynamic arm motion sequence with maximum exaggeration, $\lambda = 1$. **Fourth Row:** Additive noise motion sequence with sinusoidal noise of ±0.15 radians, $\lambda = 1$. Images are spaced at intervals of 0.067 seconds.

The sensitivity measure, $log(\alpha)$, is defined as

$$log(\alpha) = \frac{log\big(H/(1-H)\big) - log\big(F/(1-F)\big)}{2}, \qquad (1)$$

where H is the fraction of pairs in a set that were *different* and which the subject labeled correctly, and F is the fraction of pairs in a section that were *the same* and which the subject labeled incorrectly

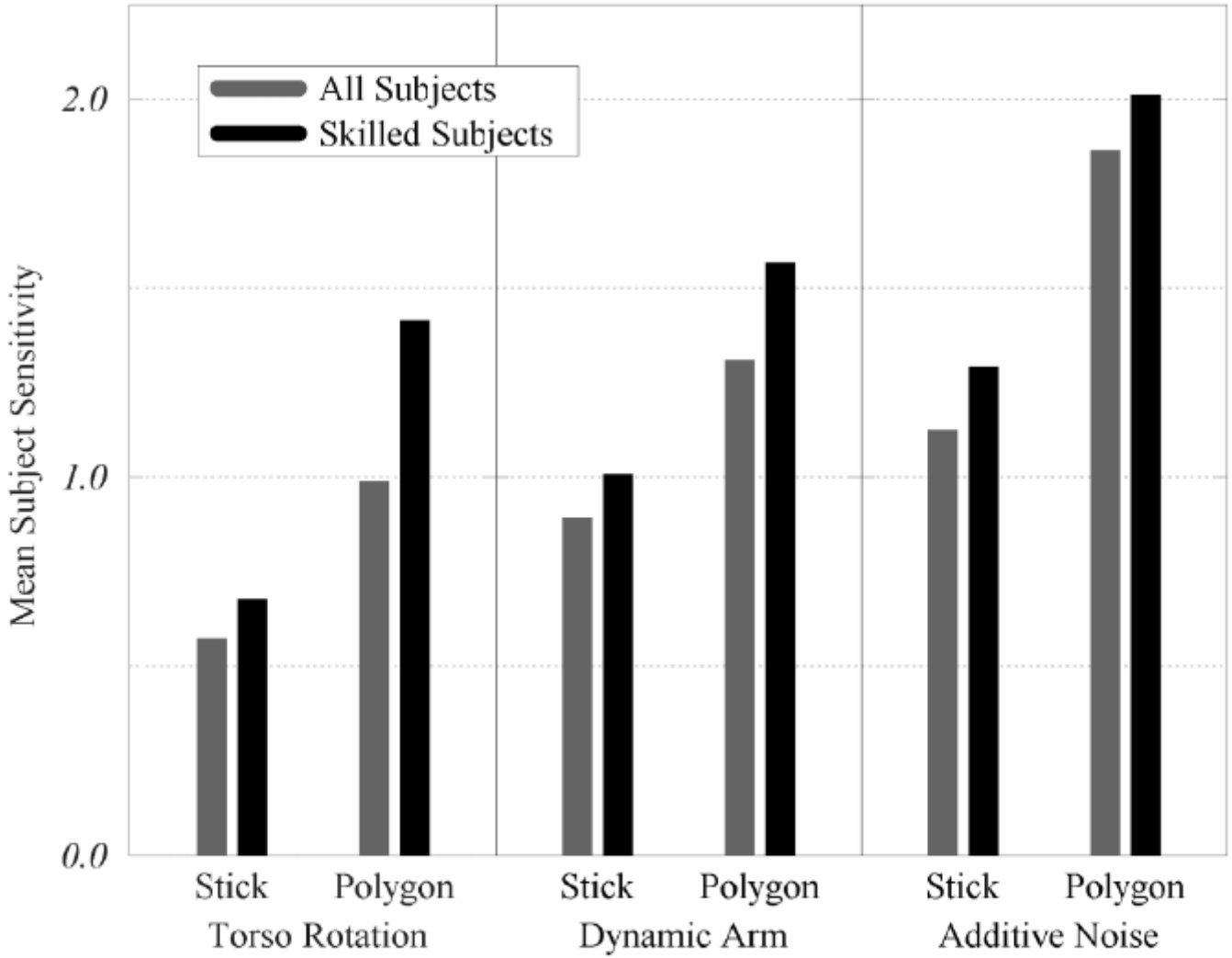


Fig. 8. Sensitivity scores by experiment averaged over subject groups. Skilled subjects are those who achieved a sensitivity score of $log(\alpha) \geq 1.0$ on either the polygonal or the stick figure portion of the test. Note that sensitivity scores are consistently higher with the polygonal model.

Conclusions

Full model allowed finer distinctions for all three of our tests.

Different models allow different distinctions to be made ➔ the graphics community should have standards for results to be compared.

Work should be produced in a way that is close to the rendering style for the final product.

Strengths? Weakness?

First study that looked at this question.
Confirmed several times since in similar but different experiments.

None of the running motions looked natural?
Did we span the space of variations?
Only tested two models (both fairly crude)
Subdivision into skilled and not skilled subjects (post-hoc)

Follow-on Studies?

Camera motion?
Clothing, Hair motion?
Breathing, facial expressions?