Dynamics Based 3D Skeletal Hand Tracking Stan Melax, Leonid Keselman, Sterling Orsten Intel Corporation



Take off the oven mitts for natural user interaction

Abstract

This research explores a new approach to tracking hands, or any articulated model, by using rigid body simulation. Based on a depth sensor's samples, the system generates constraints that limit motion orthogonal to the rigid body model's surface. These constraints, along with prior motion, collision/contact constraints, and joint mechanics, are resolved with a projected Gauss-Seidel solver.

Rigid Body Dynamics

Our approach to hand tracking draws its technical core from the world of ragdoll physics. We phrase pose tracking as a linear complementarity problem, and use a projected Gauss-Seidel solver to find a solution.

Each frame, we generate surface constraints based on our depth camera's samples, in addition to mechanical constraints from the hand model. We then run a small, fixed number of solver iterations. This flexible approach enables us to easily incorporate additional information and heuristics, as elaborated below.

Hand Model

We construct a model of the human hand using seventeen convex polyhedral parts (bones). These bones overlap slightly at joints, and are constrained to limit motion to physically plausible states. We use a reference hand model 20 cm long, authored in 3D Studio Max[®], and can parametrically scale different bones to track users with different sized hands. As our model is specified entirely in data, our generic tracking system is able to track other articulated objects without retraining.



Multiple Simulations

Our efficient simulation framework is extendable and flexible. Conjectures, heuristics and externally-derived information can all be formulated as additional constraints in the solver.

To test a variety of competing hypotheses, we run multiple biased simulations, each incorporating the corresponding constraints. The outcomes of the simulations are evaluated and the best fit one is returned.



©2013 Intel Corporation





Voxel Subsampling

We utilize volumetric subsampling to generate improved depth samples from the camera. It allows us to both discard sparse camera noise (by rejecting insufficiently dense voxels) and produce a point cloud with uniform spatial sampling across the surface of the tracked object.



ACMSIGGRAPH



Results & Usages

Our approach is able to track the fully articulated pose of a user's hand in a wide variety of configurations and motions. We are even able to handle the difficult case of a largely occluded hand model, where the depth information is ambiguous. For example, we are able to correctly determine which finger is bent from a side profile configuration.



By using our tracked pose information, it is possible to use a powered ragdoll approach to create an interactive physics playground. In these examples, we use the Bullet physics engine to enable interaction with blocks and swing sets without any specialized code. The level design becomes the interaction design.





References

- Vertex Number



- O Depth Samples Closest Feature
- Legend



Hand Pose: 6 DOF for each of the17 bones



Samples mapped to nearest rigid body of hand tracking model.







CATTO, E. 2005. Iterative dynamics with temporal coherence. In Game Developer Conference (GDC) 2005. 2. GANAPATHI, V., PLAGEMANN, C., KOLLER, D., AND THRUN, S. 2012. Real-time human pose tracking from range data. In ECCV 2012 3. OIKONOMIDIS, I., KYRIAZIS, N., AND ARGYROS, A. 2011. Efficient model-based 3D tracking of hand articulations using Kinect. In BMVC 2011.

> stan.k.melax@intel.com leonid.m.keselman@intel.com sterling.g.orsten@intel.com