# CS231A Project Proposal: Real-Time Airbending

Mridul Aanjaneya
Stanford University
aanjneya@stanford.edu

Michael Lentine
Stanford University
mlentine@stanford.edu

## Abstract

*We propose to design a framework for real-time interactive physics-based simulation using state of the art motion sensing devices such as the Microsoft Kinect. The user would be able to interact with the system using simple gestures, where each gesture would have a predefined action associated with it. Unlike previous work on gesture recognition, we would like our recognition to be 'action aware' but pose-invariant, meaning that the gesture should be invariant w.r.t view as well as direction. This will require novel pose descriptors which would take into account the intrinsic relationships between different body parts rather than the actual embedding in 3D space. The recognition process will also require the knowledge of a metric on the space of such descriptors for accurate matching. Moreover, the real-time interaction would require fast algorithms for both physical simulation and gesture recognition. We plan on evaluating our method by computing precision and recall for the gesture recognition algorithm and providing an interactive demo.*

## 1. Introduction

Real-time simulation has started gaining wide interest because of its potential applications to video games as well as special effects. Current state of the art algorithms for simulation attempt to approach real-time by coarsening the discretization of high-fidelity numerical methods. While it is necessary to achieve visual realism through simulations, it is also important to expand the domain of possible inputs for better interaction. With the recent development of the Microsoft Kinect, many researchers have started considering its possible applications to enhance the user experience.

## 2. Gesture Recognition

Gestures are perhaps the most natural way of interacting with a given system. While gesture recognition has been an active area of research, most available systems fail to capture the semantic content inherent in gestures and rather focus on recognizing them through extrinsic relations between different body parts [1]. This contradicts the very definition of a gesture. For example, if one wants to shoot a fireball to the right and to the left, the performed action remains the same, but the user shoots in two different directions. Current algorithms will incorrectly interpret these actions as two different gestures. In order to recognize gestures invariant w.r.t. direction and view, one needs to design novel descriptors which encode the semantic meaning of an action and intrinsic relations between various body parts, as opposed to the extrinsic 3D embedding. Moreover, since a gesture is a temporal sequence of poses, one needs to learn a metric on the space of such descriptors for accurate matching of the actual gesture. With these two techniques in place, one can now encode a gesture as a temporal sequence of descriptors and match two gestures using the dynamic programming edit-distance algorithm for string matching. If time permits, we also wish to explore the problem of sampling such temporal sequences for encoding a gesture compactly while still preserving its semantic meaning.

## 3. Real-time Interactivity

While gesture recognition is necessary for an interactive simulation, real-time algorithms for physically based simulations also need to be developed. In order to achieve plausible results, many current methods rely on fine discretizations in both space and time, making these techniques impractical for real-time applications [2]. To alleviate these problems, we plan on implementing methods that reduce the cost of individual steps in a traditional simulation algorithm as well as techniques for conserving physical quantities and thus allowing for accurate coarser discretizations.

## References

[1] C. Chen, Y. Zhuang, F. Nie, Y. Yang, F. Wu, and J. Xiao. Learning a 3d human pose distance metric from geometric pose descriptor. *IEEE TVCG*, 17(11):1676 – 1689, 2011.

[2] R. Fedkiw, J. Stam, and H. Jensen. Visual simulation of smoke. In *SIGGRAPH 2001*, 2001.