

# Gesture – based Sound Source Localization and Manipulation

Shashank Ranjan

University of Florida, Department of Computer and Information Science and Engineering

## Introduction

Gestures as mode of human-computer interaction have gained popularity in recent years with advancement in computer vision and depth sensor technologies. Usage of sensors such as Microsoft Kinect in gaming applications have helped increase the general awareness of these devices as well augment the comfort level of their use. Meanwhile, 3D audio applications continue to suffer from lack of intuitive mode of user interactions. Using this system, we show that gestures can be a natural and intuitive alternative method of localizing sound in 3D audio applications.

## Objective

Manipulate sound location using gestures

## Target Population

Researchers, Composers, Artists, Sound Engineers

## Examples of Existing Solutions

- Air Violin
- V – motion project
- AHNE (Audio – Haptic Navigation Feedback) - Grabbing Invisible Sounds with Magical Gloves : Open Gestures, But with Sound and Feel Feedback

## Problems with existing solutions

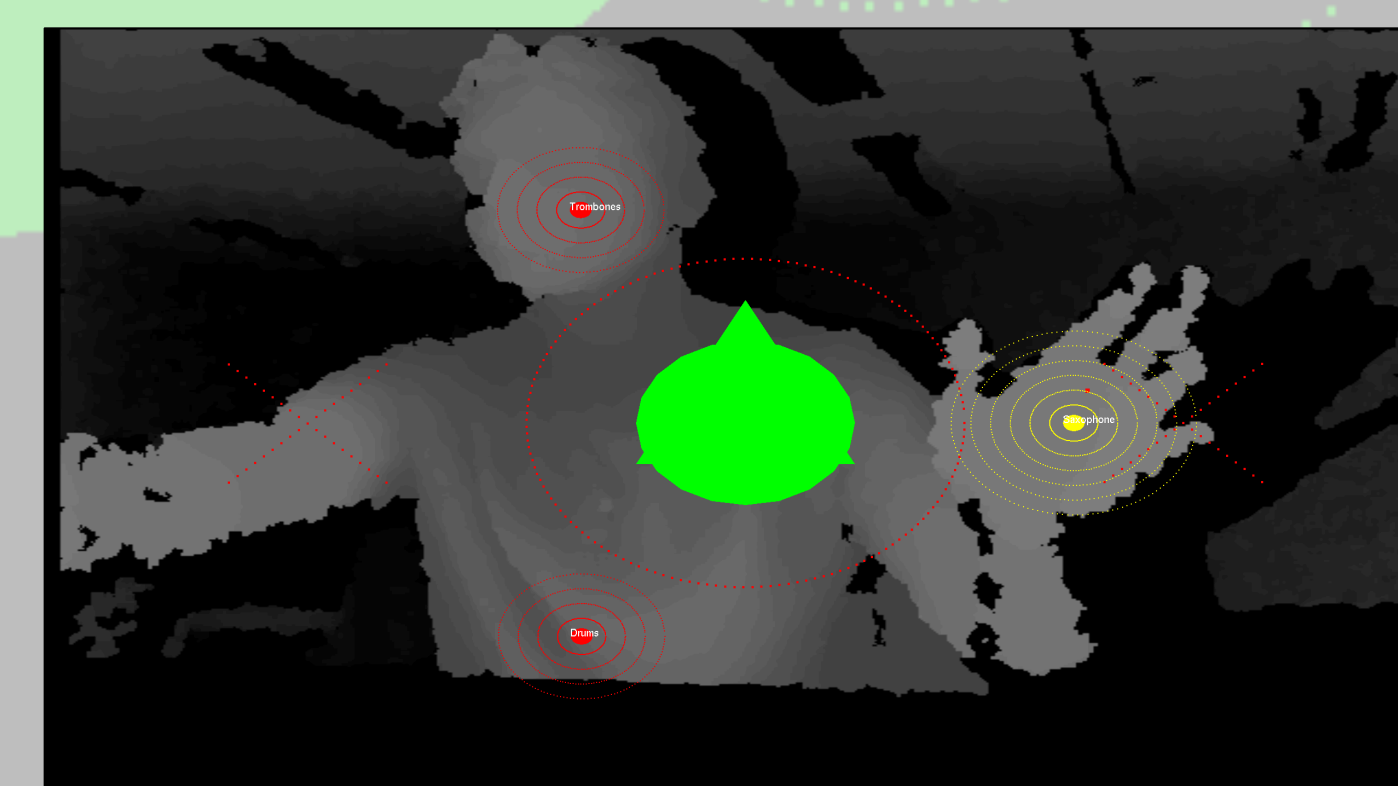
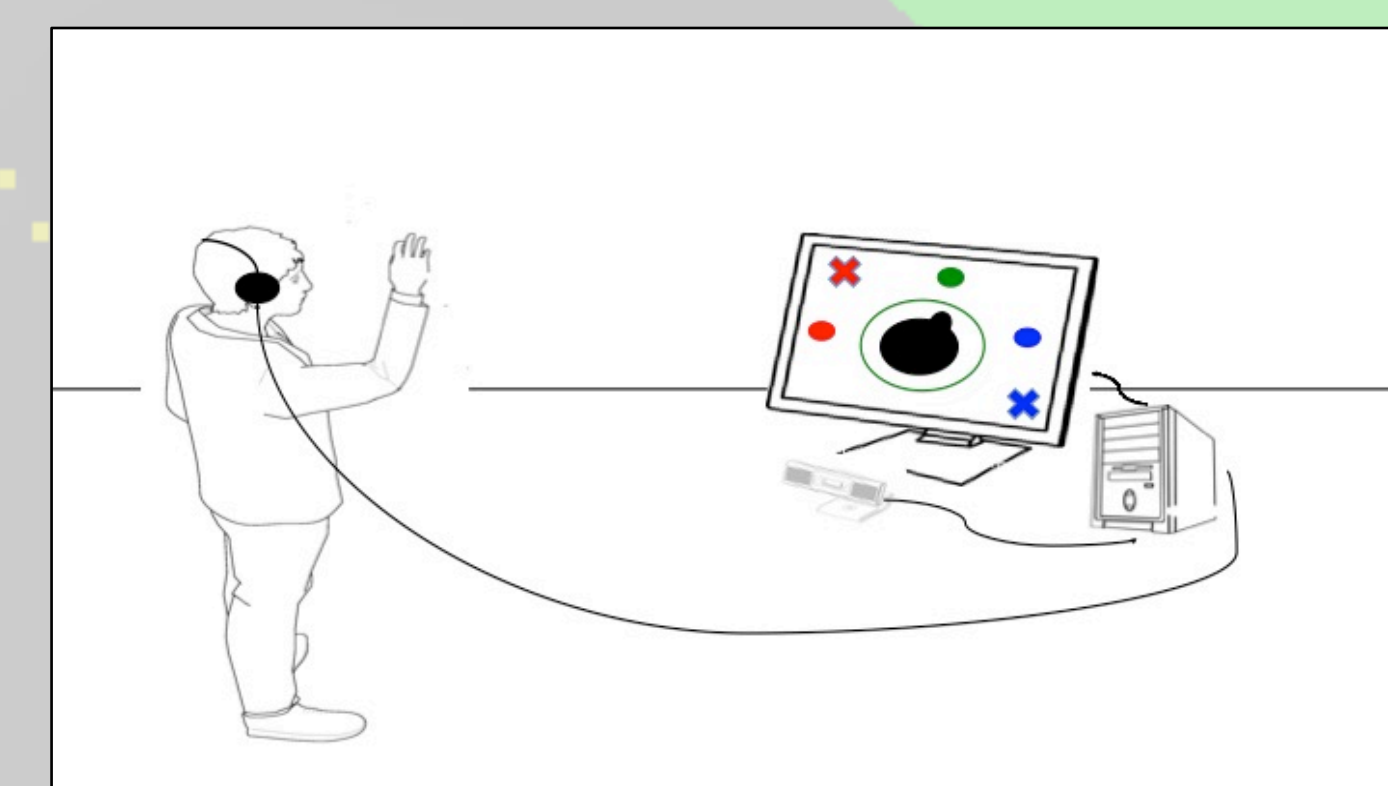
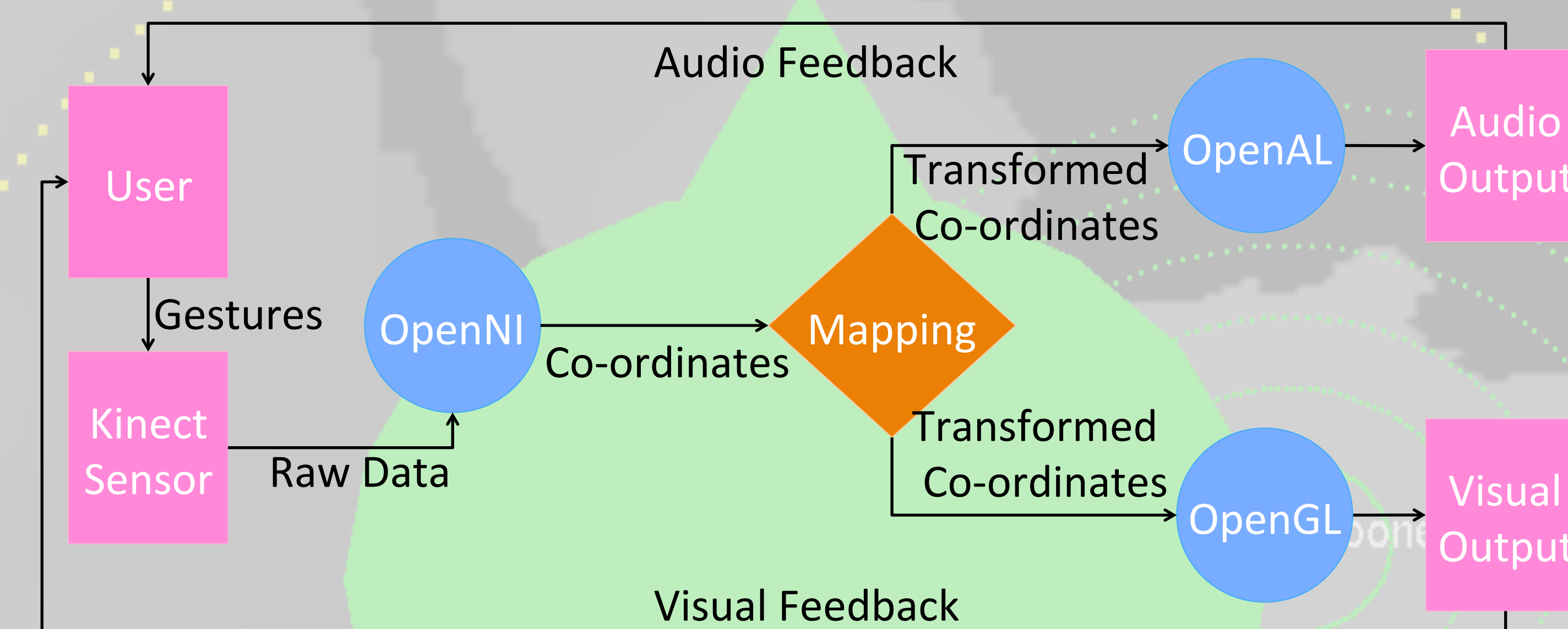
- Concentrate on sound creation
- Require specialized and expensive hardware

## Design

Hardware:

- Microsoft Kinect's sensor
- Headphones
- Monitor or a display
- CPU

Software:



## Gestures

Command	Gesture
To start	Wave
Foucs	Move hand on top of the source
Select	Push/Grab
Release	Push/ Release

## Testing the System

Stimuli : Drums, Guitar and Saxophone

Setup : 3 sources, 3 targets

Procedure:

1. Move two sources to the designated target positions
2. Move a source along a path
3. Measure time taken for each task
4. Measure accuracy
5. Take feedback on the ease of use of the system

## Initial Findings

- The tasks were easy to understand
- Selection and placement of sound sources was difficult
- User was able to move the sources easily
- But movement along a fixed path was difficult
- User could recognize and distinguish between different sound sources
- User could perceive the sound source moving
- Some errors in pinpointing the exact location of source

## Future Work

- Better gestures for selecting and moving sources
- User Head Tracking – move the listener as well as source
- Add and remove sound sources dynamically
- Changing elevation of sound – currently only azimuth is changing
- Make the visual interface 3 dimensional

## Acknowledgement

I would like to thank Dr. Kyla McMullen, Assistant Professor, CISE at University of Florida for guiding this project.