

אוניברסיטת בן-גוריון בנגב

Ben-Gurion University of the Negev

הפקולטה למדעי ההנדסה

המחלקה להנדסת חשמל ומחשבים

Faculty of Engineering Science

Dept. of Electrical and Computer Engineering

פרויקט מסכם בקורס מבוא עיבוד תמונה

Final project in the course introduction to image processing

Preliminary report

Air drums virtual system using 3D stereo reconstruction

|  |  |  |
| --- | --- | --- |
| מספר הפרויקט: | p-2018-049 | Project number: |

|  |  |  |
| --- | --- | --- |
| סטודנטים  (שם ו ת.ז.): | Roy Hirsch 305052920  302724893 Ori Chayoot | Students  (name & ID): |

|  |  |  |
| --- | --- | --- |
| מנחים: | Dr. Tammy Riklin Raviv | Supervisors: |

|  |  |  |
| --- | --- | --- |
| תאריך הגשה: | 2017 | Submission date: |

1. Abstract:

Our project’s goal is to create a virtual air drums system using object detection and 3D reconstruction. The user will be recorded “playing” with drum’s sticks in the air, the system will estimate the spatial location of the sticks and will produce sounds as the user “hits” the virtual drums. Our system will include a static setup which will hold a couple of RGB camera above the user. The two cameras will record and analyze the users’ movements in real-time.

1. Project’s goals:

* Spatial estimation of the sticks location, we will estimate the sticks location over two planes:

XY plane – we will estimate the location of the sticks over the different drums in the set.

Z plane – we will estimate the distance of the sticks from the cameras plane in order to determine if there was a hit.

* Real-time calculation, in order to be able to respond quickly to the user’s movements our algorithm must be efficient and precise.

1. Description of the system
   1. Planned algorithm

The first step of the algorithm will be calibrating the two cameras before every record session. This is a critical step in 3D reconstruction. After the cameras are calibrated we will be ready to record a session with the user.

A record session will contain two processing steams – a stereo stream and a mono stream.

The mono stream will use one cameras output to detect the sticks and the stereo stream will use the two cameras to build the depth image.

The mono stream:

* Sample the video stream, crop the relevant ROI of the virtual drums set.
* Pre-processing of the stream may include rectifying the image, bias and gradient corrections etc.
* Object detection – will use feature extraction and
* The output of this process will be binary map of the detected object. The map will be transformed to the 3D reconstruction block.

The stereo stream:

* Sample the video stream, crop the relevant ROI of the virtual drums set.
* Pre-processing of the stream may include rectifying the image, bias and gradient corrections etc.
* 3D reconstruction of the object’s area. In order to speed up our process we will use the prior of the object’s mask (extracted in the previous process) in order to calculate the disparity map of a smaller region. We will use the camera calibration and the object’s mask to triangulate two images’ small ROI and will produce a disparity map.

The output of the 3D reconstruction will be a map of the detected object with its distance from the cameras plane.

The decision maker will process this map and will determine the spatial location of the sticks. It will determine if there was a “hit” (distance is under specific threshold) and which drum should play. The decisions will be transferred to a virtual instrument which will play the relevant sound.

* 1. Block diagram

Two RGB cameras

one RGB cameras

Pre-processing

Pre-processing

3D reconstruction

Object detection

Cameras calibration

Object’s mask

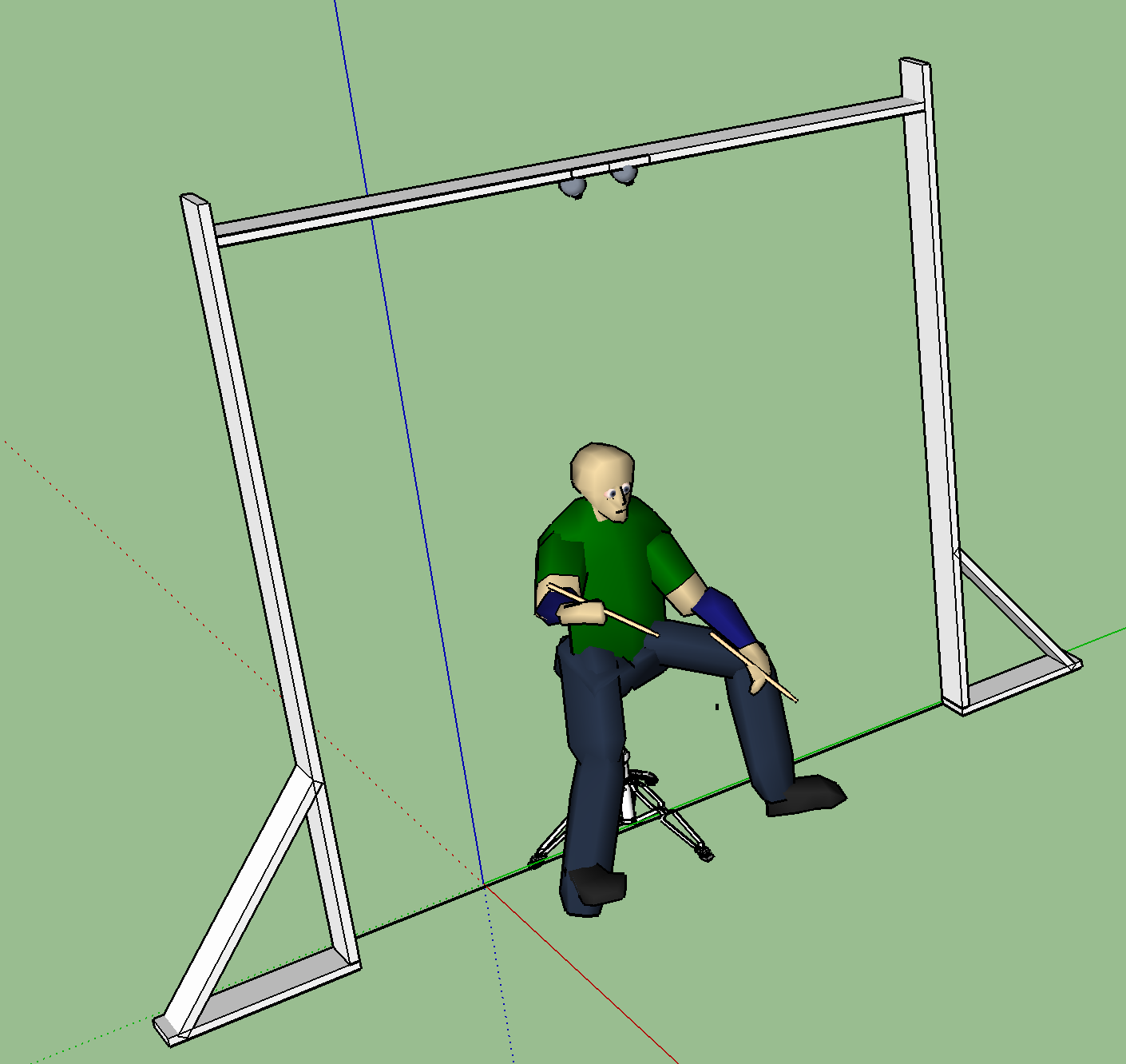
Decision maker

Virtual instrument

RGB & depth point cloud

* 1. System illustration

This is an illustration of the static setup we would like to build, where the two cameras are located above the over.



1. References: