

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 ות.ז.): | Roy Hirsch 305052920  30**2724893**  Ori Chayoot  Asaf Manor 201605110  Gal Lifshitz 301780441 | Students  (name & ID): |

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

|  |  |  |
| --- | --- | --- |
| תאריך הגשה: | February 2018 | Submission date: |

1. Abstract:

Our project’s goal is creating a virtual drumming system using object detection and 3D reconstruction which functions as follows. At first the user will be recorded playing with drum’s sticks in the air. Next, using object detection and 3D reconstruction our system will estimate the spatial location of the sticks in 3D coordinates and produce sound which simulates the user playing actual 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 drum sticks location:

XY plane – estimation of the sticks location using methods of segmentation and temporal estimation.

Z plane – estimation of the sticks distance from the camera’s plane 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, and take advantage of prior information acquired from the temporal sequence of images.

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 camera’s output to detect the sticks XY location, and the stereo stream will use the two cameras to build the depth image.

The mono stream:

* Sample the video stream, crop the relevant Region of Interest (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 color masking.
* The output of this process will be binary map of the detected object and a calculated Center of Mass for each stick. The map will be transferred to the 3D reconstruction block.
* The calculated Center of Mass can be used as prior information to estimate the next ROI in the image for calculation efficiency,

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. 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, given the spatial location, will determine the required action. 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. We maintain the option to supply other parameters to the Decision Maker such as velocity in the different axes of the drum sticks, and even approximated acceleration.

* 1. Block diagram

Two RGB cameras

RGB camera

Pre-processing

Pre-processing

3D reconstruction

Object detection

Cameras calibration

Object’s mask

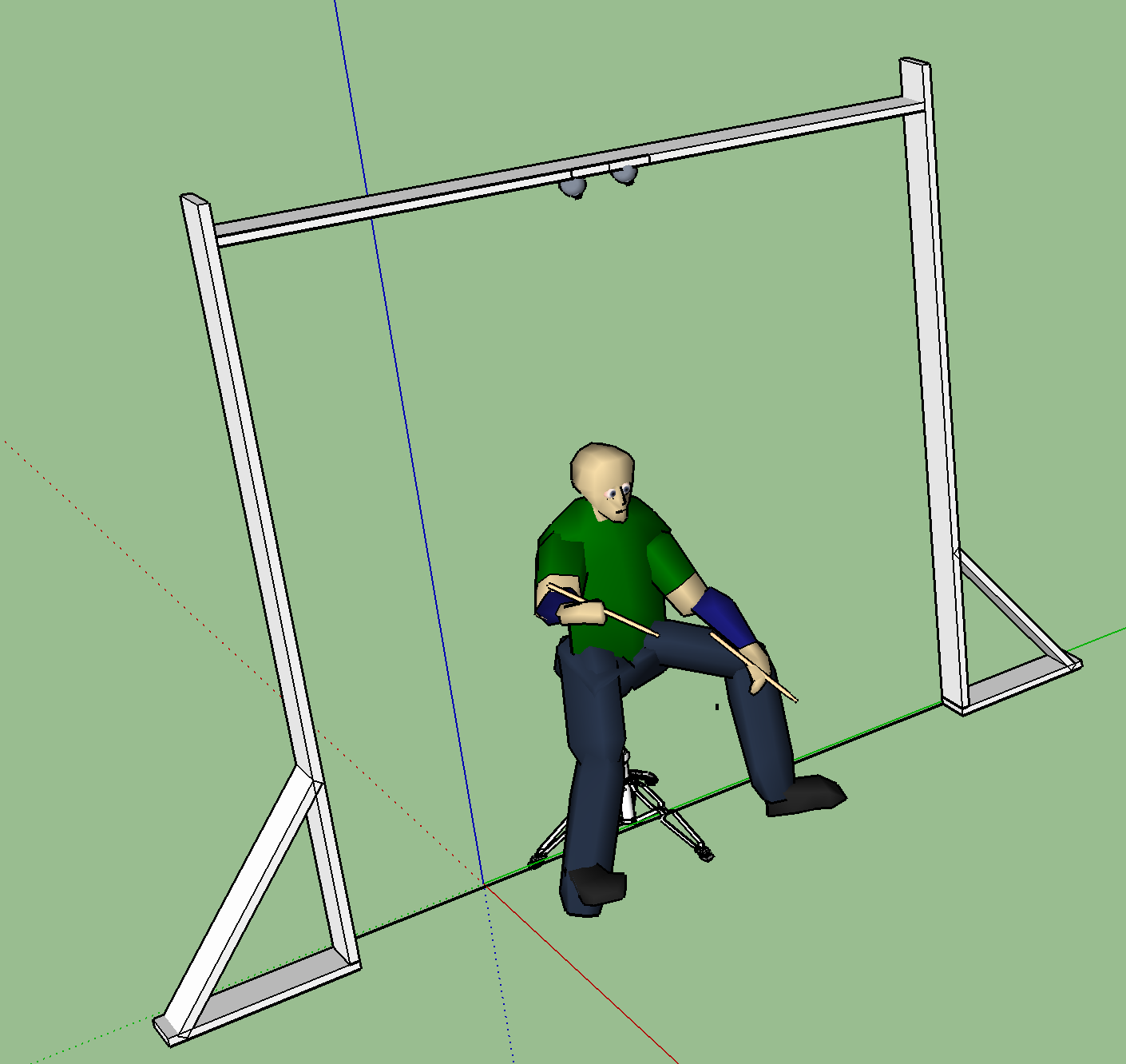
Decision maker

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:
   1. Li Tang, Chengke Wu, H. T. Tsui and Shigang Liu, "Algorithm for 3D reconstruction with both visible and missing data," in Electronics Letters, vol. 39, no. 23, pp. 1640-2-, 13 Nov. 2003.
   2. Y. M. Mustafah, R. Noor, H. Hasbi and A. W. Azma, "Stereo vision images processing for real-time object distance and size measurements," *2012 International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, 2012, pp. 659-663.
   3. J. Bang, D. Kim and H. Eom, "Motion Object and Regional Detection Method Using Block-Based Background Difference Video Frames," *2012 IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, Seoul, 2012, pp. 350-357.
   4. M. Xu and Z. Yu, "3D image segmentation based on feature-sensitive and adaptive tetrahedral meshes," *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, 2016, pp. 854-858.