

V-Drum: An Augmented Reality Drum Kit

Nestor Lobo¹

Scholar, Electronics Department, Vivekanand Education Society's Institute of Technology, Mumbai, India¹

Abstract: Most modern drum kits are large in size and difficult to transport. Similarly, the cost of a fully stocked drum kit is far in excess of what an average enthusiast can afford. The maintenance of such equipment can be expensive and keeping the instruments properly tuned involves a considerable amount of time. All this can discourage many from seriously learning to play percussion instruments. Musicians draw inspiration from the world around them and in many cases are unable to construct a new piece due to the lack of readily available instruments. The V-Drum is a low-cost, portable drum kit whose implementation is based in Augmented Reality. With the V-Drum, any person with interest can learn to play percussion instruments and even construct music with the variety of virtual instruments at their disposal.

Keywords: Human Computer Interaction, Image Processing, Musical Instrument, Augmented Reality, Pattern Recognition, Computer Vision.

I. INTRODUCTION

A. V-Drum

A drummer, as a musician, is an integral part of any band. It is the drummer who provides the band with the beat and timing which is vital in the creation and performance of music. A drummer, however, requires a large amount of space given the sheer size and number of instruments. Each instrument also requires its own stand. This makes it difficult to move a drum set, since it requires a secure means of transport, lest any instrument get damaged.

Most musicians compose music at irregular intervals, often drawing inspiration from their interactions with the environment around them. However to actualize such inspiration they may not have easy access to a studio, or the ability to carry their drum set around with them.

Like all musical instruments, percussion instruments requires regular tuning. The instrument's level of tuning varies every time the instrument is played or transported. The quality of the reproduction of sound is affected by climatic factors and so must be removed from and placed back into cases every time the instrument is to be used. All this contributes to the wear and tear of the instruments.

The V-Drum is offered as a solution to such problems. It is a Human Computer Interactive system that provides a virtual drum kit at the user's disposal. The system is an application of Computer Vision and Augmented Reality. Fig. 1 depicts a basic representation of the V-Drum.



Fig. 1. System Representation

Given the absence of the need to physically move instruments about, the V-Drum can be used in any environment without any worry about the sound quality of the instruments being affected by environmental factors. A user can shift location easily without having to carry along accessories for every percussion instrument. It is low in resource costs making it an ideal learning environment and easily supports the regimen of regular practice.

B. Analysis of Physical Components of Percussive Sound Production

The user has 3 items at their disposal to interact with the system, 2 Drum Sticks and a Drum Pedal.

A drum stick has 3 sections of interest as shown in Fig. 2; the cylindrical shaft which is the area where the user can grip the drum sticks, the shoulder is the tapering section just above the shaft and the tip or bead is used to strike the instrument. The pair of drum sticks are held one in each hand and is moved about the plane of the set.

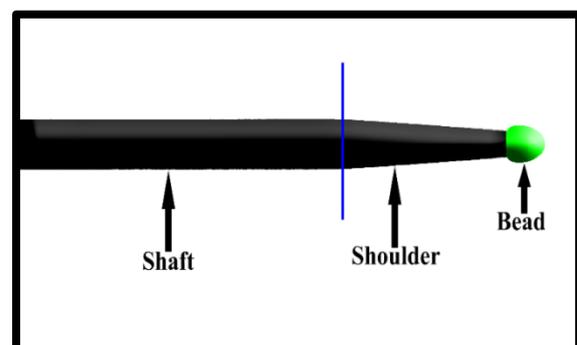


Fig. 2. Sections of a Drum Stick

The Drum Pedal shown in Fig. 3, consists of a footplate that is depressed to move a drive mechanism downwards, thus swinging the mallet forward into the drumhead. When the footplate is released, the mallet draws back and returns to its initial position. The drum pedal thus executes a single plane movement into the drumhead.

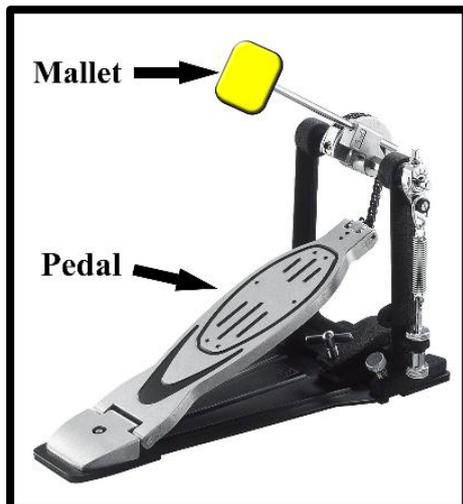


Fig. 3. Drum Pedal Sections

The beads and the mallet are the objects of interest to V-Drum system and are searched for in the camera input as described later.

II. PROPOSED WORK

The user starts the V-Drum and sits within the view of the camera. Images of a selection of the available instruments are displayed on the screen. The image of the user is obtained. This image is thresholded to search for the drum stick beads. If a bead is detected, the bead is tracked. If the beads move over any of the available instruments, the corresponding sound is played. This sound is modulated by the velocity of impact. The user can add or remove any of the available instruments. The user can also move the instruments or change their sizes. A basic block diagram of the implementation is shown in Fig. 4.

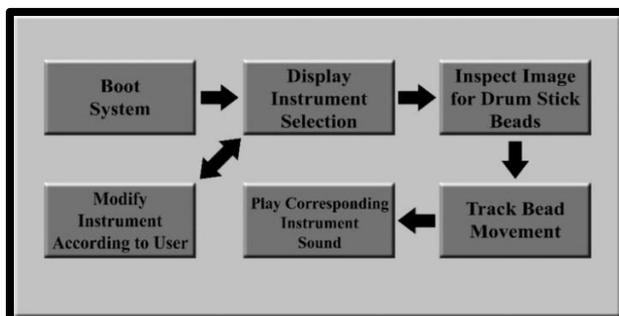


Fig. 4. System Block Diagram

III. METHODOLOGY

A. Hardware

The V-Drum requires a camera for acquiring the image of the user. This can be a standard USB camera having a Frames per Second rate in excess of 15. The higher the frame rate, the better the system becomes. The other hardware that influences performance is the processor. The V-Drum can be implemented on a Raspberry Pi model 2 B having 1 GB of RAM and a 900 MHz Processor. The V-Drum requires an input system such as a mouse and keyboard for user interaction but the system can also function if the display is a touch screen monitor.

Additionally, the V-Drum can also be implemented on a user's laptop or personal computer.

B. Initial Setup

After boot, the camera takes an image. On this image, images of the drum kit selection are displayed as per the last saved user settings or the default arrangement. The kit instruments are initialized with information such as their name, location, size, and corresponding sound in the system. The volume, set to the default, is also displayed. Users may adjust the volume as per their requirements. The metronome can be started and the user can choose to begin recording their interactions with the system. The initial system is as shown in Fig. 5.



Fig. 5. Initial System display.

C. Image Processing

The Image Processing is accomplished with the Python programming language using the OpenCV Libraries. The Numpy Libraries are also used for performing matrix operations.

The image that is obtained from the camera is in the BGR (Blue, Green, and Red) format. This image is filtered to discard noise. The image is transformed in the HSI (Hue, Saturation and Intensity) color format. This aids in filtering. This image is filtered to include only the chosen colors of the drum stick beads, and for the kicker mallet as shown in Fig. 6.

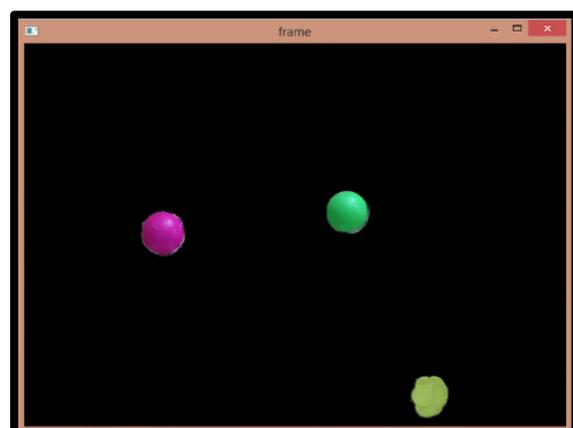


Fig. 6. Background is filtered out, leaving the beads. Red and Green: Drum Stick Beads, Yellow: Kicker Mallet.

This image is then checked for the profiles of the beads. If a satisfactory match is made, the location and size of the beads are determined. Using this data, the beads are highlighted as shown in Fig. 7.



Fig. 7. Drum stick Beads detected.

D. Criteria for Accepting Input from Drum Sticks

Once the beads have been detected, the Euclidean Distance of the beads from the drum kit selections on screen is calculated, and if it lies within the range that corresponds to the instrument sizes, (let us term this as an instrument's 'active area'), this is considered to be a contact between the bead and the instrument and the corresponding instrument is played. The image is also highlighted with a randomized color to visually depict contact as in Fig. 8.

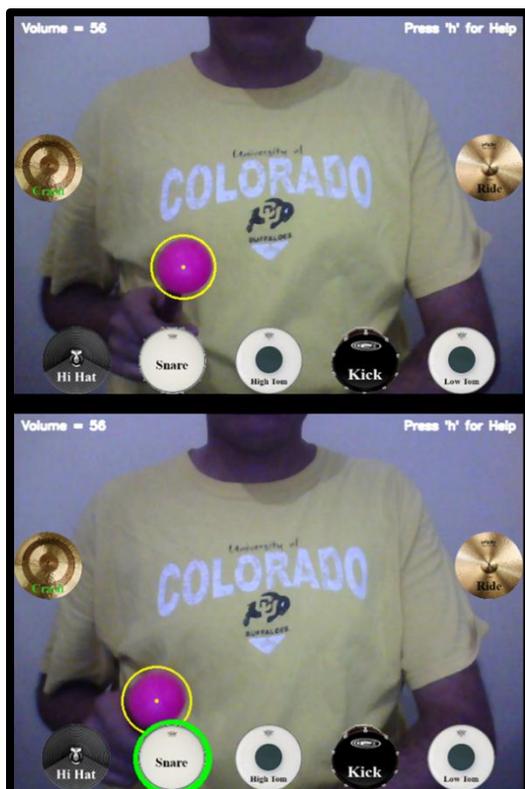


Fig. 8. Top: Pre-Impact Display, Bottom: Post-Impact Display.

If the instrument is struck again before the sound fades out, the sound is stopped and restarted, as a typical instrument would respond. Once the instrument is struck, the bead must leave its *active area* completely, else the sound will not be played again as in Fig. 9. This is a crucial step or else the instrument will repeatedly play its sound. Only when the bead leaves the *active area* first and then returns, will the response be accepted as an input and the sound played.

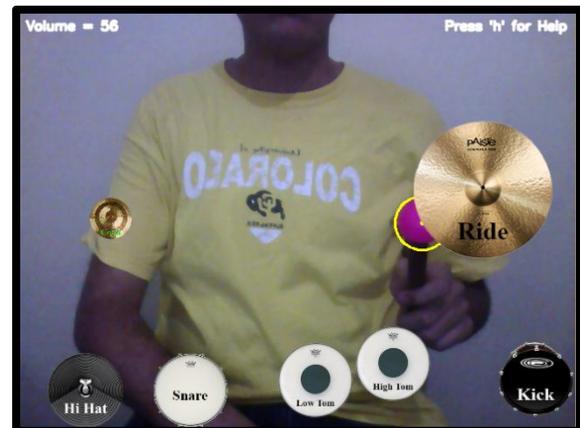


Fig. 9. Input not recorded after the initial impact, system waits for the bead to be withdrawn from the instrument.

E. Kickstand Input

A drummer typically utilizes 3 forms of interaction with the drum kit. They are the two drum sticks and an additional kicker for the bass drum. The base drum kicker here is given its own detectable bead. Whenever the users depresses the pedal, the mechanism sends mallet upwards and into the camera's view. When the mallet is detected, this is taken as kick input and the kick drum is sounded.

F. Instrument Properties

Each instrument is initialized with information such as its name and sound file. This is used when an instrument is struck or the user is recording a session. Information about the instrument's size and location are extracted from the configuration file.

These properties are innate to each instrument allowing it to be treated as separate entity when interacting with the instrument as in the change of location and size. These interactions are affected using a location input from a hardware mouse or a touch screen.

G. Change of Location

Users can move instruments around at any point, so long as they are within the constraints of the screen. If the user is a left handed person, they may move the entire kit around to suit their preference.

When an instrument is moved, its Euclidean Distance is calculated anew and properties like the *active area* of the instrument are modified according to the new location which make the instrument directly playable as shown in Fig. 10.

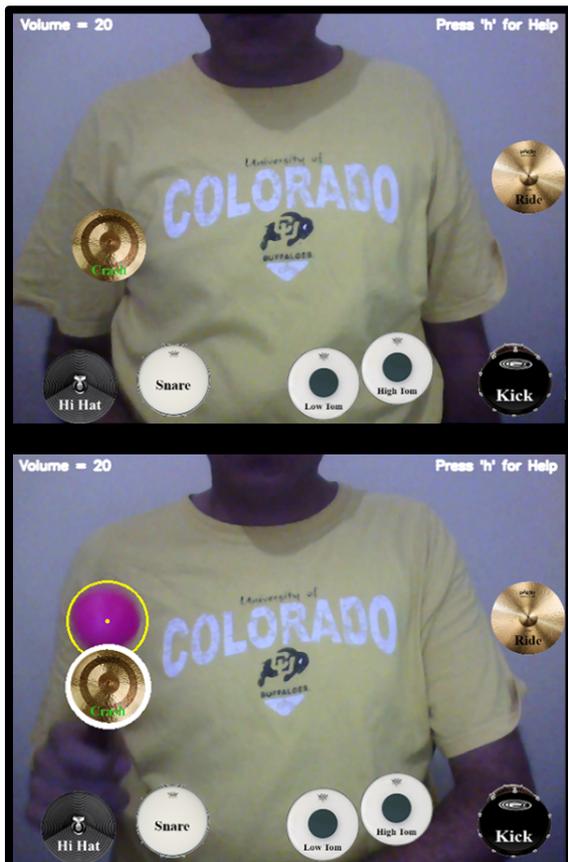


Fig. 10. Top: Instruments are moved. Bottom: Moved instruments are directly playable.

H. Size Modification

Instruments can also be resized as per the user's need as shown in Fig. 11. The *active area* of the instrument increases radially while keeping the location constant. A user who uses a specific instrument more can size it larger. As the user adds more instruments to the kit, they can resize them all to fit within the screen reducing the limit on the number of instruments the user can have on screen at a time.

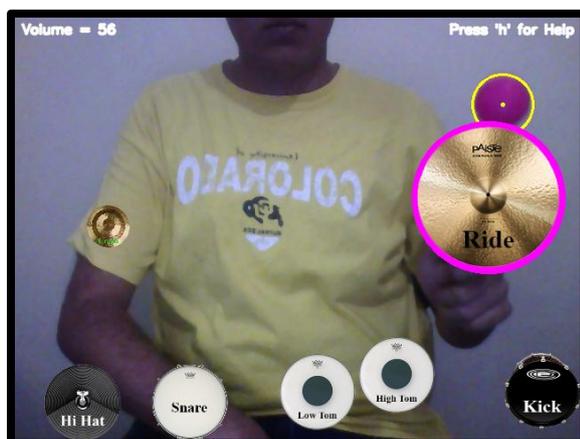


Fig. 11. Resized instrument is played.

I. Kit Configuration

The V-Drum allows users to switch between kit items on the go. Users can add any percussion instrument in the

system by calling up the drop down menu. The selected percussion instrument then appears on the screen and the user can move and size it as required. The flow of this system is shown in Fig. 12. Users can also delete items at any point and these items are moved to the top of the stack of instruments making them easy to retrieve if the user needs them again. A user can also choose to completely change the drum kit to a different style i.e. from a Rock kit to a Metal kit. In this case, the user must only select a change of styles and the entire kit displayed on the screen is replaced by the new kit.

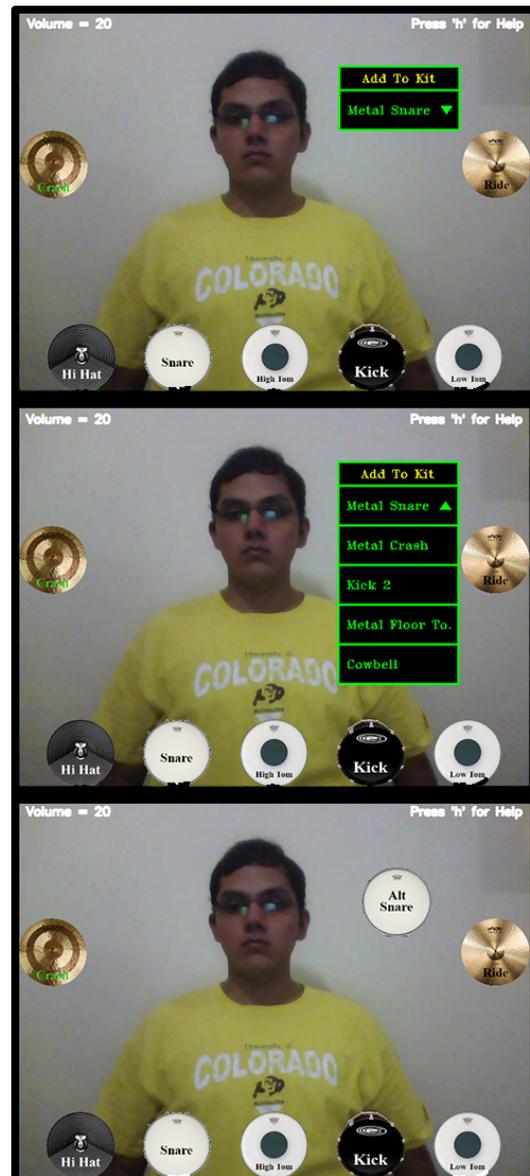


Fig. 12. Process of adding a new instrument to kit (Top to Bottom).

This new kit has instruments and tuning that corresponds to the chosen style setting. When a user exits the system, a configuration file is saved with information as to the locations and sizing of instruments that the user has made. This configuration file is loaded at the start so that the user is saved time rearranging their kit. This can also be restored to the default settings.

J. Velocity Controlled Modulation

An important feature of a physical drum kit, that the sound produced depends on the strength of the strike, is replicated in the V-Drum system. The velocity of the drum stick bead is calculated using system time as a clock source and the distance is calculated from the Euclidean Distance between its previous position and last position. This velocity is used to modulate sound output. Thus a faster hit would produce a louder sound and a hit of lower velocity would produce a softer sound output.

K. Metronome

A metronome is a device that produces rhythmic ticks. The rate of these ticks can be modified thus providing a variable tempo. The V-Drum has its own metronome system that produces ticks as per the tempo set by the user which aids in keeping to a beat. The metronome runs parallel to the image processing and so neither functionality is interrupted.

L. Recording

The user can use the V-Drum system to record the session. When a user selects record, a file is created that stores information as to which instrument was played, how loudly it was played, and the time gap in between consecutive inputs. This information can then be converted by the system to an mp3 or wav file as the user chooses. The user can thus record sessions of their playing. The system described is more efficient, as compared to a constant recording, as it only needs to take into account the user's interactions with the instruments and the time gap between them. A video of the user's session can also be recorded.

IV. WORKING

The user faces the camera and boots the V-Drum system. The user is greeted by a start screen detailing the basic working of the system and information on how to modify the volume or move and resize instruments. Once the user proceeds from this screen, the camera acquires an image and displays the instruments as per the configuration file. The user may rearrange the instruments, change the volume or start the metronome. The camera input is examined for the drum stick beads and kicker mallet. The user can then play the percussion instruments by moving the drum sticks and using the drum pedal. The user's session may be recorded for viewing and editing later.

V. FUTURE SCOPE AND APPLICATIONS

The V-Drum system can be used as a teaching aid for music schools or a home user interested in learning the basics of various percussion instruments. It can also be an excellent tool for regular practice for musicians, especially off-site practice sessions. Giving the ease of switching between instruments, a user could easily learn the sound of a new instrument and learn how to adapt it into their musical style. The instruments are all ready to play and the pitch of the instruments are good and do not fade in the V-Drum system. All these are very crucial properties for students who have just started learning a new instrument.

The system can be used for instruments like a xylophone. The keys of the xylophone would be arranged in order and a user would play music in the same manner as in the percussion instruments of the V-Drum. For instruments like the Tabla or Conga, a user's gloves could be used as the input system instead of the drum sticks of the V-Drum. The user can also add special effects to their kit and use them like an instrument.

VI. CONCLUSION

The implementation of the V-Drum system was successfully carried out by the author. Its low resource cost, ease of modification of instrument selection, and portability make it an attractive system for musicians who can use it on-the-go to learn, and create music. Its various additional functionalities such as the metronome, and the recording facility add to the user's experience and create a well-rounded system.

REFERENCES

- [1] G. Bradski, "The OpenCV Library" Dr. Dobb's Journal of Software Tools (2000).
- [2] Stéfan van der Walt, S. Chris Colbert and Gaël Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation", *Computing in Science & Engineering*, 13, 22-30 (2011).
- [3] Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing* (3rd Edition) (31 August 2007).
- [4] (2011) Pete Shinnars, "PyGame". [Online]. Available: <http://www.pygame.org/>
- [5] Lenz, Reimar K. and Roger Y. Tsai. "Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology" *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10, 5 (September 1988), 713- 720.
- [6] Kim, Won S. "Advanced Teleoperation, Graphics Aids, and Application to Time Delay Environments" *Proceedings of the 1st Industrial Virtual Reality Show and Conference (IVR '93)* (Makuhari Meese, Japan, 23-25 June 1993), 202-207.
- [7] R F Badness, *Drum Programming: A Complete Guide to Program and Think Like a Drummer* (1 November 1991).
- [8] P Magadini, *Drummer's Guide to Music Theory: The Drummer's Resource to the Rudiments of Music* (2nd Edition) (1 December 1986).

BIOGRAPHY



Nestor Lobo is currently pursuing his B.E. in Electronics Engineering at V.E.S. Institute of Technology, Mumbai. His proficiencies include Embedded Systems, System Conception and Design, Image Processing and Machine Vision.