# Ontology Based Recognition of Hasta Mudras in Bharatanatyam Dance

Aryan Agarwal (19CS30005)

Under supervision of Prof. Partha Pratim Das

April 16, 2022

# Contents

# 1 Introduction

## 1.1 What is Bharatanatyam?

Bharatnatyam is a form of Indian Classical Dance that originated from the southern India. Bharatnatyam traditionally comprises of a solo dancer who is adorned in a saree and jewelleries. The dancer performs various body and hand gestures (hastha mudras) which convey different meanings.
Hastha mudras are of 3 types:

- Single Hand Gestures
- Two Hand Gestures
- Dance Hand Gestures

Out of these, single hand gestures (there are 32 of them) are relatively easy to represent using Ontology. The dancer conveys the *ras* (sentiment) and *bhava* (mood) of the story through their hand gestures and facial expressions.


Figure 1: Hastha Mudra Examples

## 1.2 Why Ontology of Hastha Mudras?

Computer vision is a fastly growing topic of research. Interaction of humans with computers is not just limited with keyboards and mouse. We now have AR/VR headsets and joysticks to play games, interact with each other on the metaverse or even do complex surgeries. In this project, we will explore these topics and try to recognise a Bharatanatyam hastha mudra using Ontology. We will also discuss a powerful method to recognise some body-based gestures.

## 1.3 What is Ontology?

Simply put Ontology is the 'nature of being'. Ontology is "what we mean when we say what we say". In computer science, Ontology of something is the representation, formal naming, and definition of the categories, properties, and relations between the concepts, data, and entities.

## 1.4 Related Work

**An Ontology for Reasoning on Body-based Gestures** [1]: Jean Vanderdonckt et al. made an ontology for structuring body-based gestures based on the user body and body parts, gestures and environment and encoded it in OWL. They also conducted an GES (gesture elicitation study) with 24 participants and classified the gestures according to the ontology.

# 2 Ontology For Body-Based Gestures

An ontology was designed to represent the body-based gestures. The user, sensor and the physical environment were expressed in the Ontology Web Language using RDF (Resource Description Framework) triples. The format of triples was <subject,predicate,object>

For example:

- "Dhoni plays cricket"
- "Right Hand is above Head"

Some main classes are:

- User : Info about who the user is.
  Ex. Body, BodyPart, Arm, Limb, Leg, Bone, Joint
- Sensor : raw data provided by a sensor.
  Ex. Kinect, VicoVR
- Detected gestures and poses : Gesture, GestureSegment

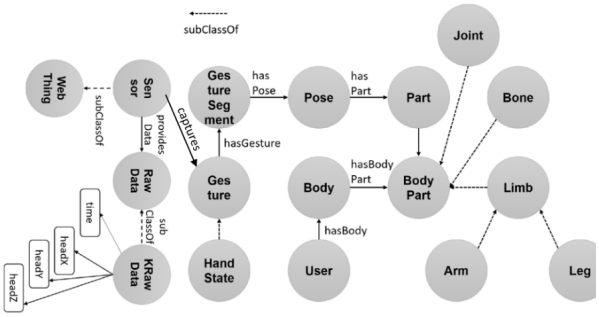


Figure 2: The ontology for body-based gestures

## 2.1 Gesture

A gesture is made up of strokes that are binding points. And by connecting these points we can define a gesture. Quite similar when we make a stroke from a paint brush by applying it in one direction. A sensor captures the gesture as tuple of 5 elements, $p_i = (x_i, x_i, z_i, w_i, t_i)$, where the last element represents the time and the first 4 are the 4D coordinates of each gesture point. Joints are defined as objects which are compared wrt their relative positioning in Cartesian space. These help in expressing the higher level relationships like Left before Right, Left meets Right, Left overlaps Right, etc. Joint are then merged into segments which gives a bone like structure and a human continuum from hands to feet.

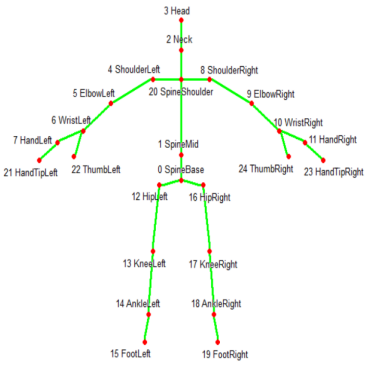


Figure 3: The 25 joints of a MS Kinect Skeleton

## 2.2 Pose

Just a comparison between two joints. Logical constructs for combining poses are:

- Merging two poses
  - Operation - and
  - Symbol - &
- Matching at least one pose
  - Operation - or
  - Symbol - |
- Matching the opposite of pose

- Operation - not
- Symbol - !

Shortest distance between the joints is calculated for maintaining proportion between various differences like difference in height, difference in angle etc. Limbs are considered as the connecting bones between hips and ankle and, shoulder and wrist and are used for calculation of rotation and angles. Degree of confidence is calculated from comparing the position of the body found (which is obtained from a Kinect frame) with a pre made model of the pose.

### 2.3 Sensor

The Sensor class is used to identify the user. It has property of 'providesData' for specifying the relation between sensor and the raw data. Raw data is expressed by another class called KRawData. KRawData stores the numerical position of various body parts such as the head, neck, hand etc.

Example:

```
:kinect rdf:type Sensor ;
    providesData (:data1, :data2,..., :dataN).
:data1 rdf:type KRawData ;
    time 0.0110013^^xsd:double ;
    headX 0.0021632^^xsd:double ;
    headY 0.8330318^^xsd:double ;
    headZ 2.2899450^^xsd:double ;
    ...
    headLeftX -0.2059579^^xsd:double ;
```

### 2.4 Other classes

- User:
  Populated by a subset of properties coming from the Person class
- Body:
  Has multiple instances of BodyPart class
  Identified with 'trackingId' functional property.
- Body Part:
  Super class of Limb, Bone, Joint etc.

## 3 Conducting the Gesture Elicitation Study (GES)

Referents were selected which are common in IoT and were divided into 3 groups:

- Unary action
  Ex: Start a player
- Binary action
  Ex: Answer/end a call
- Linear action
  Ex: Increase/decrease volume

24 voluntary participants were recruited for the GES. They had different occupations and reported frequent use of computers and smartphones. Experiment was conducted in a laboratory. Sheets of paper having the referent written on them were provided to the participants and the participants were recorded in front of a camera and a Kinect camera.

### 3.1 Procedure

Participants were asked to sign a consent form, and were provided information about the process of the experiment. They were also asked about to fill a questionnaire about their sociodemographic status and use of technologies. The referents were randomly shuffled before giving to each participant for which they elicited the gestures. After the end of each session, participants were again asked to fill some questionnaire about the study.

## 4 Results

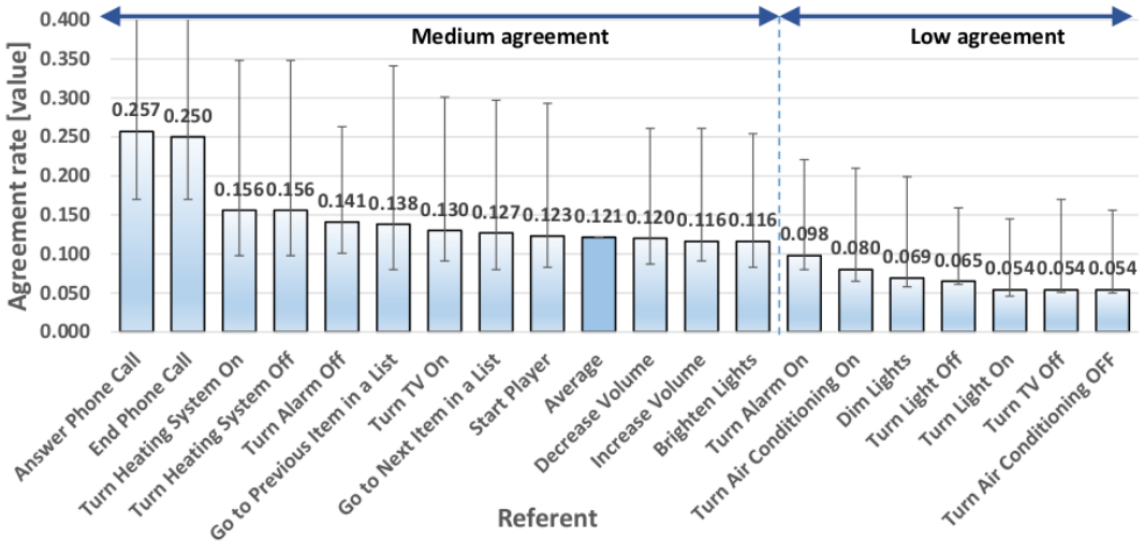


Figure 4: Referents in decreasing order of their agreement rate. Error bars show a confidence interval for $\alpha = 0.05$

The figure depicts the referents in decreasing order of their agreement rate computed by AGaTE. Overall, agreement scores and rates are medium in average magnitude, in particular for rates (which are the most demanding ones) between 0.257 and 0.116 for the global sampling. Results fall inside medium consensus category with their average in the same interval.

## 5 Conclusion

We discussed an ontology for structuring body-based gestures based on the user, bodypart, environment etc. which was encoded in OWL. Ontology based recognition of Hastha mudras in Bharatanatyam dance can draw a lot of ideas from this experiment's gesture set. Ontology provides extensibility and thus the Ontology of Bharatnatyam dance can also be further enhanced for recognition of other dance forms.

## 6 Future Work

My future work will be to make a Hastha Mudra retrieval system using Ontology. The input will be search query for the Hastha Mudra and my software will then give all the images or videos containing the queried hastha mudra from the corpus. For example, we can search for all the videos or images containing the *Pataka* hastha mudra.

## References

[1] M. Ousmer, J. Vanderdonckt, and S. Buraga, in Proceedings of the acm sigchi symposium on engineering interactive computing systems (2019), pp. 1–6.