* 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 2 types:-
  + Single Hand Gestures
  + Double Hand Gestures
* The dancer conveys the *ras*(sentiment) and *bhava*(mood) of the story through their hand gestures and facial expressions.

**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.

**Why Ontology of Hasta Mudra**

* In my BTP project, we will create a Bharatanatyam Hastha mudra ontology and try to recognise those Bharatanatyam hastha mudra using Ontology.
* We can represent the structural and semantical information of Bharatnatyam using Ontology, which can help us extend the Ontology to other dance gestures in future.
* 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.

**Related Work**

* An Ontology for Reasoning on Body-based Gestures.
  + 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

Ontology For Body-Based Gestures

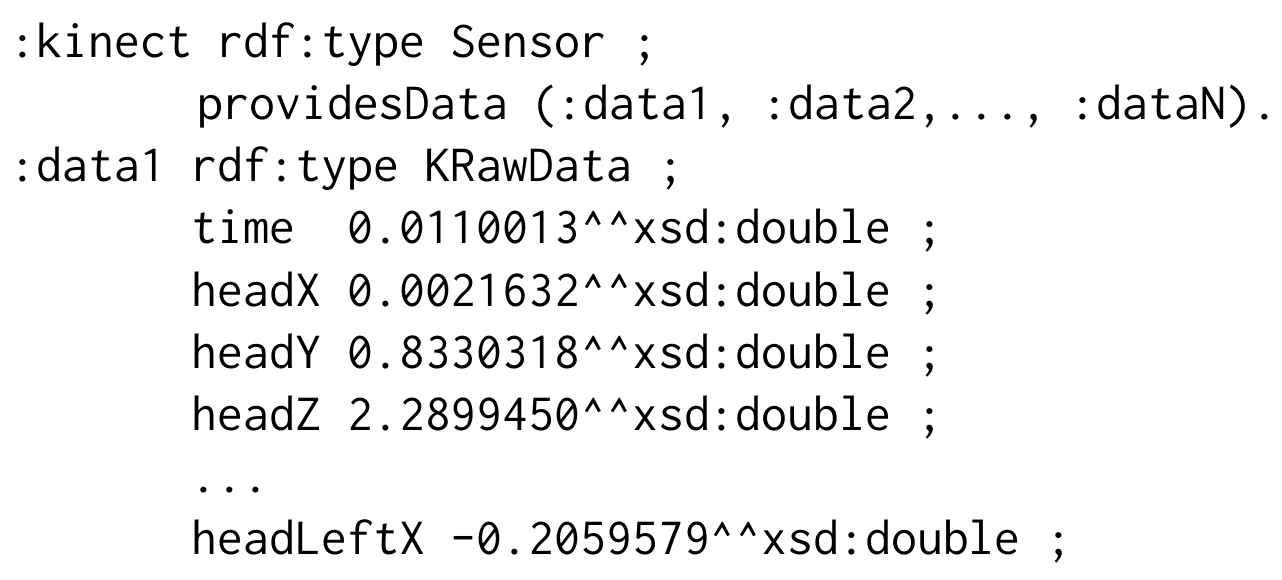
* An ontology was designed to represent the body-based gestures. The user, sensor and the physical environment were expressed in the Ontology Web Language using RDF (Resource Description Framework) triples.
* The format of triples was <subject,predicate,object>
  + “Dhoni plays cricket”
  + “RightHand isAbove Head”
* Some main classes:-
  + 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
* **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, pi=(xi,yi,zi,wi,ti) , 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.

**Pose**

* Just a comparison between two joints. Like difference in angles, distance etc.
* Logical constructs for combining poses :-
  + Merging two poses.
    - Operation - and
    - Symbol - &
  + Matching at least one pose
    - Operation - or
    - Symbol - |
  + Matching the opposite of pose
    - Operation - not
    - Symbol - !
* Since The distance between the wrist and the hand can be smaller than between hand and elbow, the difference in height and angle between the two should be proportional, resulting in the same confidence rate when comparing both. For maintaining this proportion, we needed to calculate the shortest distance between any two given joints and add the positions of each intermediary joint to the final ratio.
* 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.

**Sensor**

* The Sensor class is used to identify the user.
* 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 -



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.
* BodyPart
  + Super class of Limb, Bone, Joint etc.

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.

**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 before the test.
* The referents were randomly shuffled before giving to each participant for which they elicited the gestures.
* Thinking time was recorded for each referent and participant pair which is the time taken by a referent to know which gesture to perform.
* Participants were also asked to rate goodness-of-fit from 1 to 10 to express to what extent they thought their gesture was appropriate to the presented referent
* After the end of each session, participants were again asked to fill some questionnaires about the study.

**RESULT**

* The previous figure depicts the referents in decreasing order of their agreement rate computed by AGaTE algorithm.
* Overall, agreement scores and rates are medium in average magnitude, in particular for rates (which are the most demanding ones) between .257 and .116 for the global sampling.
* Results fall inside medium consensus category with their average in the same interval.

**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.