

Thesis Report 14 : 1 June - 8 June

Goals

- Extract Bandai Dataset LMA Features ✓
- Relabel Bandai Dataset Mocap ✓
- LMA-PAD using Bandai Dataset ✓
- PAD-LMA using Bandai Dataset ✓
- Start writing Thesis Document ✓

Last Week Leftovers:

Done

- **Extract Bandai Dataset LMA Features:**

- Redefined our LMA features (added 2 more and removed 1). We now have a total of 25 LMA features. We're also extracting each 5 frames rather than every 0.5 seconds for better integration with the Motion Synthesis.
- Set currently looks like this:

LMA Data format:

```
{
```

```
  frame_counter: frame at which LMA features were computed,
```

```
  label: PAD Emotional Coordinates (3D)
```

```
  lma_features: [
```

```
    # Position Features
```

```
    max hand_distance (1D),
```

```
    average l_hand_hip_distance (1D),
```

```
    average r_hand_hip_distance (1D),
```

```
    max stride length (distance between left and right foot) (1D),
```

```
    average l_hand_chest_distance (1D),
```

```
    average r_hand_chest_distance (1D),
```

```
    average l_elbow_hip_distance (1D),
```

```
    average r_elbow_hip_distance (1D),
```

```
    average chest_pelvis_distance (1D),
```

```
    average neck_chest_distance (1D),
```

```
    average total_body_volume (1D)
```

```
    average lower_body_volume (1D) (NEW)
```

```
    average upper_body_volume (1D) (NEW)
```

```
    triangle area between hands and neck (1D)
```

```
triangle area between feet and root (1D)
```

```
# Movement Features:
```

```
l_hand speed (1D)
r_hand speed (1D)
l_foot_speed (1D)
r_foot_speed (1D)
neck speed (1D)

l_hand acceleration magnitude (1D)
r_hand acceleration magnitude (1D)
l_foot acceleration magnitude (1D)
r_foot acceleration magnitude (1D)
neck acceleration magnitude (1D)
```

```
]
```

```
}
```

• Relabel Bandai Dataset Mocap:

- Analyzed the mocap from the Bandai Dataset
- Their out of the box labels don't directly translate to emotions since they have things like: "giant", "childish", "old", "masculine" and so on
- By looking at the animation previews - <https://github.com/BandaiNamcoResearchInc/Bandai-Namco-Research-Motiondataset> - created the following mapping between mocap category and emotion. Unfortunately, we don't have as many "unique" emotions. We do still, however, have at least one emotion on each quadrant regardless.
- Moreover we have several levels of the same emotion (for example we have 4 "tired/exhausted" emotions).
- **Do note that, this labeling was done by analyzing a couple of each animations of each category, but this does not mean that these coordinates are accurate to every single animation in that category. This is a possible point of failure, but it was also a problem with our previous dataset**

Bandai Category	Emotion	Coordinates
Normal (Bandai 1/2)	Neutral	(0.05, -0.4, 0.0)
Tired (Bandai 1)	Tired	(0.15, -0.7, -0.2)
Old (Bandai 1)	Exhausted	(-0.1, -0.55, -0.15)
Exhausted (Bandai 2)	Tired 2	(0.2, -0.75, -0.3)
Elderly (Bandai 2)	Exhausted 2	(-0.1, -0.55, -0.15)
Angry (Bandai 1)	Angry	(-0.5, 0.8, 0.9)

Bandai Category	Emotion	Coordinates
Happy (Bandai 1)	Happy	(0.8, 0.5, 0.15)
Youthful (Bandai 2)	Happy 2	(0.6, 0.4, 0.1)
Sad (Bandai 1)	Sad	(-0.6, -0.4, -0.3)
Proud (Bandai 1)	Proud	(0.4, 0.2, 0.3)
Giant (Bandai 1)	Confident	(0.3, 0.3, 0.9)
Masculine (Bandai 1)	Confident 2	(0.25, 0.15, 0.3)
Masculinity (Bandai 1)	Confident 3	(0.25, 0.15, 0.3)
Masculine (Bandai 2)	Confident 4	(0.3, 0.4, 0.6)
Not Confident (Bandai 1)	Afraid	(-0.6, 0.7, -0.8)
Active (Bandai 1/2)	Active	(0.1, 0.6, 0.4)

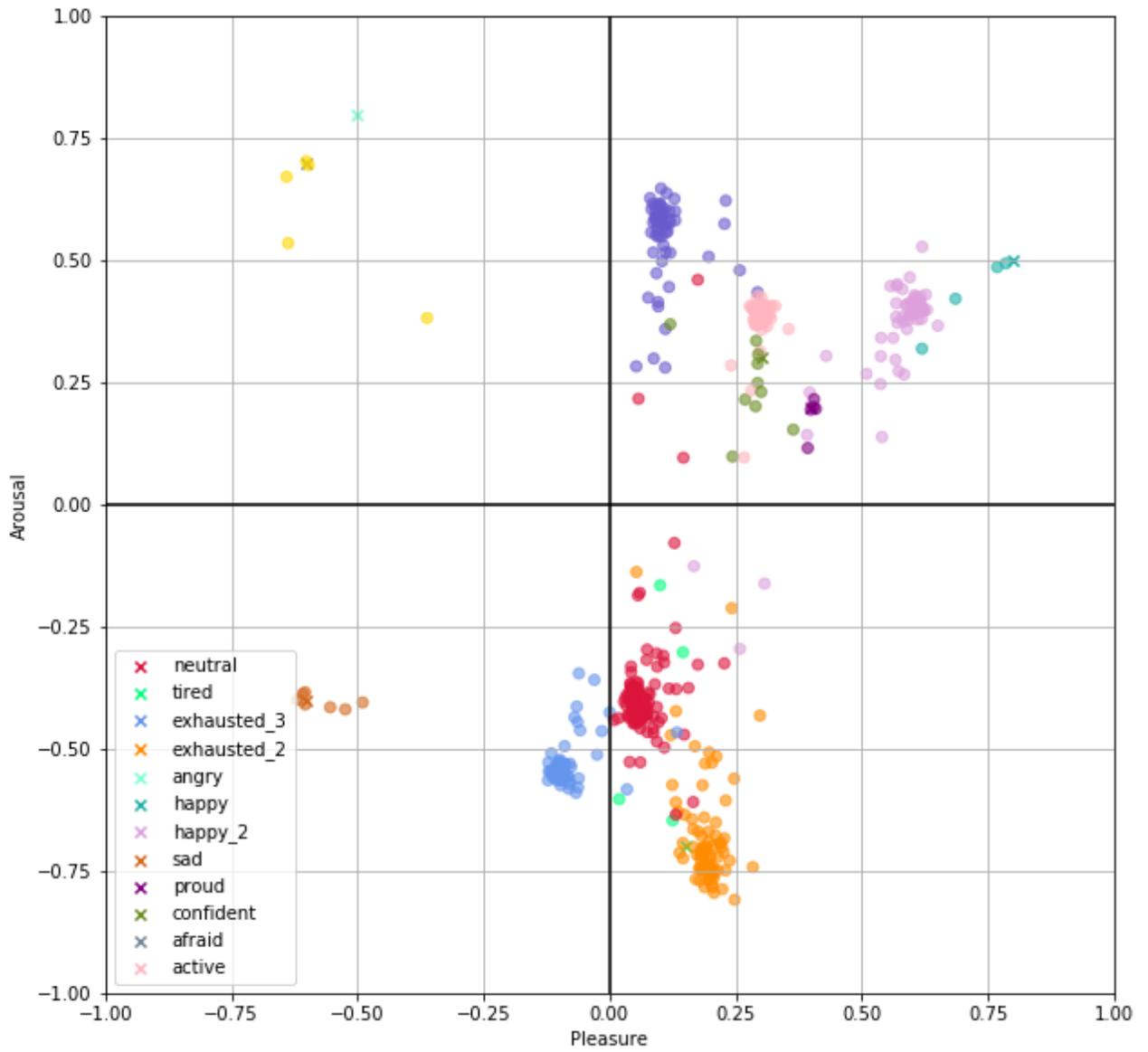
- **LMA-PAD Models:**

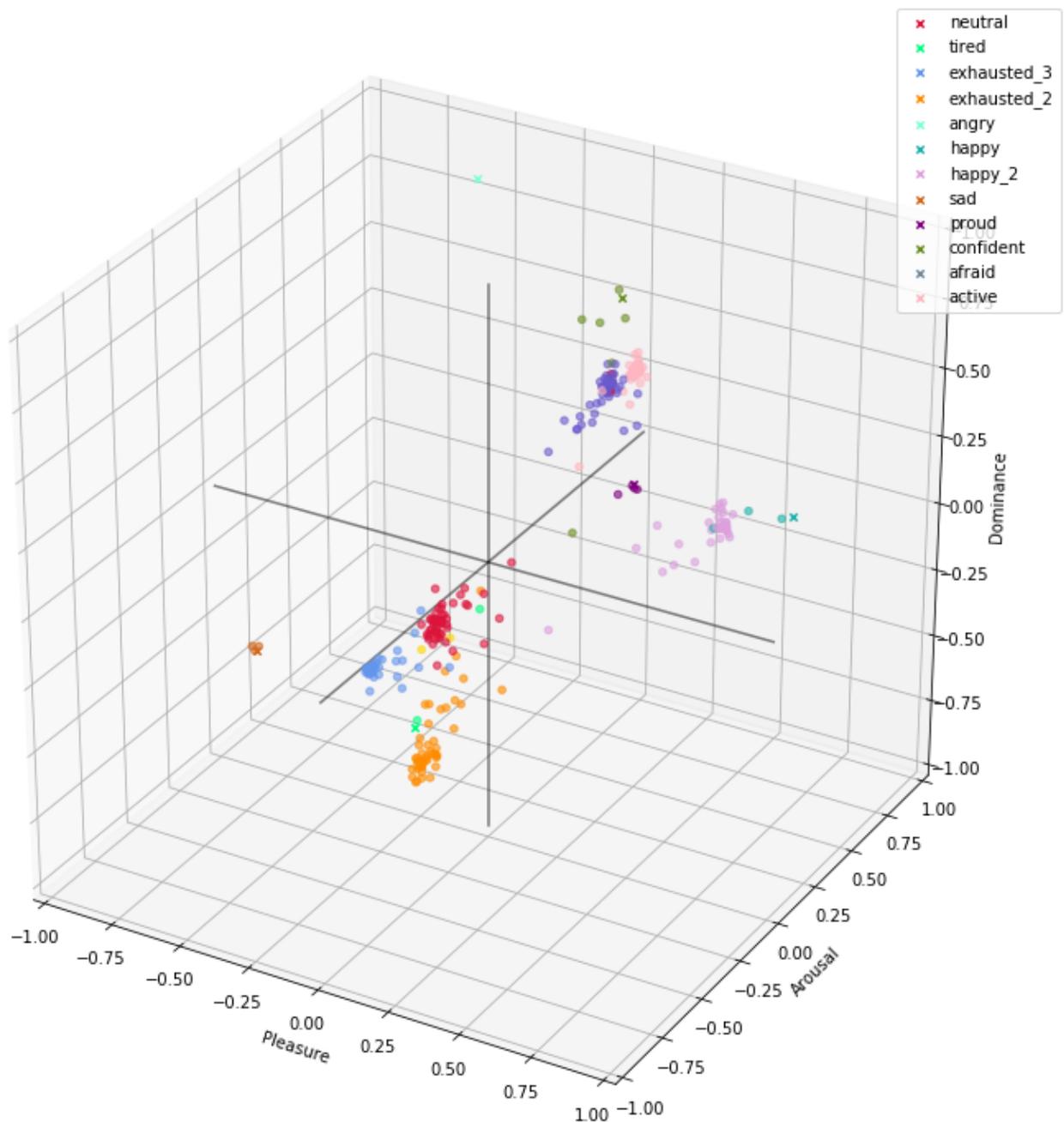
Pleasure
MSE: 0.00
MAE: 0.03

Arousal
MSE: 0.01
MAE: 0.05

Dominance
MSE: 0.01
MAE: 0.04

- Retrained our LMA-PAD models using our Bandai Dataset





• PAD-LMA Models:

- Retrained our PAD-LMA models using our Bandai Dataset
- Retried using the Autoencoder to reduce the number of features and then using XGBoostRegression to map PAD coordinates to the latent features
- **Results were better this time around. The Autoencoder is able to accurately encode and decode the 25 LMA features down to just 5 Latent ones**
- **Furthermore, XGB is able of going from 3 PAD features to each of the 5 Latent features more or less accurately** (i.e the predicted emotional coordinates of the generated LMA features achieved from going from the original PAD coordinates to Latent, and then using the autoencoder to decode them don't stray too far from the original emotional coordinates)

- **However, this only applies to PAD coordinates similar to those in our training/test dataset (i.e emotional coordinates similar to the ones we have data for).** When we specify coordinates that are different from our set, the models struggle, with the predicted emotion not being too similar to the intended one.
 - For example. We can provide our models the coordinates "(-0.5, 0.8, 0.9)" and it will properly generate a set of LMA features that properly depict these coordinates (according to our emotional prediction model)
 - However, if we specify something like "(1.0, 0.0, 0.0)", the generated LMA set is not as good. This problem comes from the fact that we never supplied the algorithms any datapoint that represented these coordinates.
- Also left XGBoost Regression models that map directly from PAD to each of the 25 LMA features training (using randomsearchCV).
- Also, the datasets we used had their LMA features extracted each 5 frames rather than every 0.5 seconds for better integration with the Motion Synthesis algorithm. Might be worth it to try using the 0.5sec interval to see if it better "generalizes the features"
- **Started Writing Thesis:**
 - Started writing thesis by transcribing the first chapters from the "Projeto de Tese" document to the intended format.
 - <https://www.overleaf.com/9859713481hfjstwkxcjx>

Left Undone

- Retrain models using 0.5sec instead of 5frame LMA extraction
- Check RandomSearchCV searches

Notes

Thoughts

This week went by fine. Worked mostly on retraining models using our new dataset. Results were pretty promising. Specifically I was pretty happy when I managed to make it so the Autoencoder was capable of properly compressing the 25 LMA features down to just 5. LMA-PAD seems to be doing great. All mean absolute errors are just under 0.05. PAD-LMA is still not really perfect. Although the autoencoder is working nicely, I'd still like to get better results on the PAD-Latent Space mapping.

Also started writing the actual thesis, mainly the introductory chapters (Introduction, Related Work and Background).

Next week, aside from continuing to write the thesis, I wanna finalize the PAD-Latent Space / PAD-LMA models and the Heuristic Rules for the Motion Synthesis Algorithm. If all goes well, next week the project should be finalized or nearly finalized meaning that we'll be done either by the 15th or the 22nd. As such we have around 1 month and a half to write the short paper for submission and the thesis document. Also, in terms of user testing, when I proposed the Projeto de Tese, I was thinking of having some users try to evaluate the system by showing them certain animations and seeing if their predicted emotions corresponded with those that the system's and whether the generated motion synthesis managed to make the character convey an emotion identifiable by a viewer. I'll probably start working on these forms either this week or next week to then send them to testers (friends, family and GameDevTecnico members) by the 22nd the latest (so that I have all the data by the start of July)

Work Hours

- Worked everyday except Saturday from about 11am to 8pm