Proof of Concept Plan for McMaster Text to Motion Database
CS 4ZP6

Brendan Duke

Andrew Kohnen

Udip Patel

Dave Pitkanen

Jordan Viveiros

October 25, 2016

1 Risks

The significant risks of this project are split into three major sections represented by the website, database, and deep learning network. The largest risk to the project is linking the three sections together so that the website can pull information from the database, and the deep learning network can use this information to run pose estimation. Some significant risks are involved within each of these sections and will be elaborated on below:

- The website must use full text search in order to correctly return information that the user searched for.

- The database must contain all the required videos and text pairings from larger libraries like Charades.

- The deep learning network is going to use Tensorflow and requires the steep learning curve that is associated with deep learning.

- Generating the proper test cases to provide proof of correctness is difficult on large databases.

2 Datasets

Our primary mode of input data for the program will be uploading an image to the website from the user's computer. If these efforts prove fruitful we will wish to We will be largely using the "Charades" Database in order to act as input for our program.

3   Deliverables

As a proof of concept demo we will show a working website that will function as a base for running pose estimation. This is to show that the associated risks mentioned above can be overcome and fully worked out as the project matures. To utilize the pose estimation the user will either have to upload an image, or search the database from a subset of the options. The prototype will have two aspects to the demonstration with the addition of a third if full text search if it is implemented on a subset of the given data. The first aspect will be a website that can be seen at the time of writing in Figure 1. Once on the website running the pose estimation will be shown by either searching for an action or uploading a picture and displaying the position of the chin, left and right humerus, left and right radius/ulna, left and right femur, left and right Tibia/Fibula and the spine of the person found within the image.

4   Performance Metrics

There are a few key aspects to which we can measure the effectiveness of our proof of concept demonstration:

- We would like to run this website from any computer at a resonable speed

- We would like the ability to take a variety of image types

- We would like the pose search function to reasonably fast

5   Resources

- Caffe

- TensorFlow

- FFmpeg

- Sphinx

FIX THIS FOR OUR REFERENCES

References

[1] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The LaTeX Companion.* Addison-Wesley, Reading, Massachusetts, 1993.

[2] Albert Einstein. *Zur Elektrodynamik bewegter Körper.* (German) [*On the electrodynamics of moving bodies*]. Annalen der Physik, 322(10):891921, 1905.

[3] Knuth: Computers and Typesetting,
`http://www-cs-faculty.stanford.edu/~uno/abcde.html`