

Text to Motion Database

System Architecture

Brendan Duke
Andrew Kohnen
Udip Patel
David Pitkanen
Jordan Viveiros

April 8, 2017

Contents

1	Overview	1
2	Anticipated and Unlikely Changes	2
2.1	Anticipated Changes	2
2.2	Unlikely Changes	2
3	Component Decomposition	3
3.1	Web Interface	3
3.2	Media Upload	4

Revision History

Date	Version	Notes
January 5, 2017	0.0	File created

1 Overview

This document will provide a general overview of how the McMaster Text to Motion Database will be constructed. The first chapter will talk about the changes that the Text to Motion Database may undergo during the development. The second chapter will contain high level component decomposition for the web interface and the options available to a user who has registered through the web interface.

2 Anticipated and Unlikely Changes

As the Text to Motion Database continues to grow and be developed there are inevitably going to be changes in the design and structure of the overall project. The next sections aim to outline the aspects of the project that may be prone to change versus the aspects that are core to the functionality of the database.

2.1 Anticipated Changes

- **User Interface Design:** The user interface will likely be changed throughout the development process due to feedback from users and supervisors. The changes should improve the user experience and add functionality when possible in order to promote the usability of the database through the web interface.
- **HTTP Server:** The HTTP server should see some changes during development as at the time of revision 0 it is still in an early stage of development compared with other modules. It will be used to upload images and videos through the command line, which will be outlined below.
- **Deep Learning Network:** The deep learning network will see improvement through the accuracy of the pose estimation performed on uploaded images or video. This may occur by re-training the existing network by the development team, or more state of the art methods and algorithms can be applied.

2.2 Unlikely Changes

- **Image/Video Upload:** At the core of the McMaster Text to Motion Database the ability to upload video and images is one of the primary functions and will remain an option throughout the project's development.
- **Database Search:** The ability to search through the database is something that the Text to Motion Database should always provide and will evolve when the full text search is implemented.

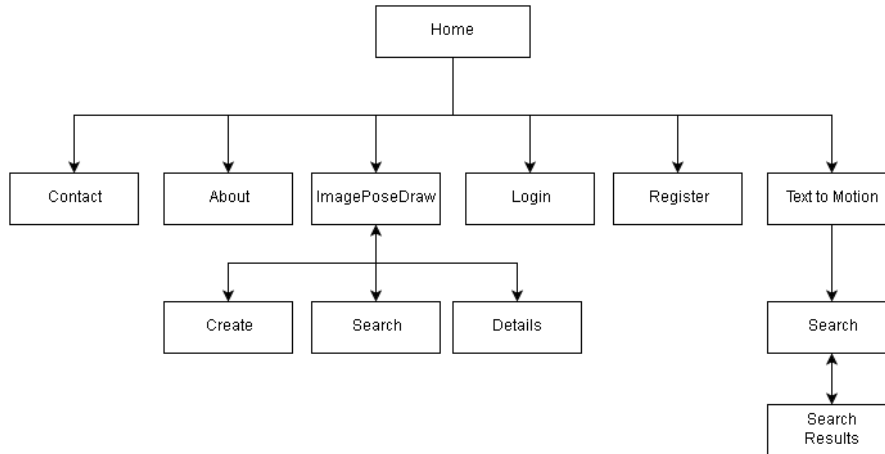
3 Component Decomposition

3.1 Web Interface

Figure 1 shows the user experience from landing on the Home page to options within the web interface.

1. Home page for the Text to Motion Database provides information and is the starting point for the journey.
2. Contact contains the contact information of the development group and its supervisors.
3. About is about the Text to Motion Database and contains a more detailed description of the application.
4. ImagePoseDraw accesses the database in order to create, display or manipulate uploads.
 - a) Create is used to take a user, who has logged in to the page in order to upload a new image or video, to the database.
 - b) Searching through the current database entries is done through the search bar located above the data table.
 - c) Details contains the description of the entry, and the options to edit, view or delete the upload.
5. Login takes the user to the login page to sign in with a registered account.
6. Register allows the user to register a new account in order to create new uploads.
7. Text to Motion provides the user the ability to utilize full text search to find associated tags.
 - a) Search is in reference to the search bar that receives input and searches the database.
 - b) Search returns the query results and displays the results to the user.

Figure 3.1: User Journey for a new user



3.2 Media Upload

Figure 2 displays the process that a verified user can go through in order to upload and interact with the Text to Motion Database.

1. The user must first be logged into the website in order to upload an image or video.
2. Using the web interface allows the user multiple options with respect to uploading images and video.
3. The standalone HTTP server allows the user to access the database remotely through the command line, or other application using HTTP requests, in order to upload images and videos and perform human pose estimation.
4. Both the web interface and HTTP server access the Tensorflow backend process in order to run pose estimation and save the results to the database for future use.
5. The database takes in search queries from the web interface and receives uploads from the pose estimation process.

Figure 3.2: A top level view of the architecture of the Text-to-Motion software suite.

