

UNIVERSITY OF WATERLOO

Faculty of Physics & Astronomy

**ACOUSTIC MODELLING USING MEL-FREQUENCY
CEPSTRAL COEFFICIENTS**

Sysomos
Toronto, Ontario

Prepared by

Thomas C. Fraser
3A Mathematical Physics
ID 20460785
January 11, 2016

154 Quarry Ave.
Renfrew, Ontario
K7V 2W4

January 11, 2016

Mr. Jeff Chen, Department Chair
Department of Physics and Astronomy
University of Waterloo
200 University Avenue West
Waterloo, Ontario
N2L 3G1

Dear Mr. Chen:

I have prepared the enclosed report “Acoustic Modelling Using Mel-Frequency Cepstral Coefficients” as my 3A Work Report for my work term spent at Sysomos in Toronto, Ontario. This is my second work term report. The purpose of this report is to examine the problems associated with a popular technique used in application design known as a spriteSheet. I aim to convince anyone familiar with the ideas discussed to consider the solution I propose to solve these problems. This report uses the techniques I developed while tackling these problems but is targeted at a audience with a wider range of applications.

Sysomos is currently working on a multi-platform video game. My supervisor, Elyot Grant, assigned me with overcoming some limitations of graphics API used.

This report was written entirely by me and has not received any previous academic credit at this or any other institution. I give permission to Sysomos to keep a copy of this report on file and use it as necessary in the future.

Sincerely,

A handwritten signature in black ink, reading "Thomas Fraser". The signature is written in a cursive style with a horizontal line above the first name.

Thomas C. Fraser
ID 20460785

Table of Contents

| | |
|--|----|
| List of Tables and Figures. | iv |
| Summary. | v |
| 1.0 Introduction | 1 |
| 2.0 Audio Survey. | 2 |
| 2.1 Social is not Clean | 2 |
| 2.2 Audio Segmentation & Environment Detection | 2 |
| 2.3 Solution Exploration | 2 |
| 2.4 Use Acoustic Modelling | 2 |
| 3.0 Proposed Pipeline. | 2 |
| 3.1 Video Downloading | 3 |
| 3.2 Audio Extraction | 3 |
| 3.3 Normalizing Signal | 3 |
| 3.4 Classification | 3 |
| 3.5 Augment Data & Frame Recommendations | 3 |
| 4.0 Mel-Frequency Cepstral Coefficients Disected | 3 |
| 4.1 Windowing | 4 |
| 4.2 Discrete Fourier Transform | 4 |
| 4.3 Triangular Windowing | 4 |
| 4.4 Logarithm of Amplitudes | 4 |
| 4.5 Discrete Cosine Transform | 4 |
| 4.6 Interpretation | 4 |
| 4.7 Deltas and Delta-Deltas | 5 |

| | | |
|-------|---------------------------------|---|
| 5.0 | Performance Evaluation..... | 5 |
| 5.1 | Binary Classification | 5 |
| 5.2 | Feature List | 5 |
| 5.3 | Feature Aggregration | 5 |
| 5.3.1 | Beat Extraction | 5 |
| 5.4 | Data Sets | 6 |
| 5.5 | Results | 6 |
| 5.5.1 | Feature Rankings | 6 |
| 5.5.2 | Classification Models | 6 |
| 5.6 | Interpretation | 6 |
| 6.0 | Conclusions..... | 6 |
| 7.0 | Recommendations..... | 7 |
| | References..... | 8 |

List of Tables and Figures

Summary

1.0 Introduction

...

2.0 Audio Survey

...

2.1 Social is not Clean

...

2.2 Audio Segmentation & Environment Detection

...

2.3 Solution Exploration

...

2.4 Use Acoustic Modelling

...

3.0 Proposed Pipeline

...

3.1 Video Downloading

...

3.2 Audio Extraction

...

3.3 Normalizing Signal

...

3.4 Classification

...

3.5 Augment Data & Frame Recommendations

...

4.0 Mel-Frequency Cepstral Coefficients Disected

...

4.1 Windowing

...

4.2 Discrete Fourier Transform

...

4.3 Triangular Windowing

...

4.4 Logarithm of Amplitudes

...

4.5 Discrete Cosine Transform

...

4.6 Interpretation

...

4.7 Deltas and Delta-Deltas

...

5.0 Performance Evaluation

...

5.1 Binary Classification

...

5.2 Feature List

...

5.3 Feature Aggregration

...

5.3.1 Beat Extraction

...

5.4 Data Sets

...

5.5 Results

...

5.5.1 Feature Rankings

...

5.5.2 Classification Models

...

5.6 Interpretation

...

6.0 Conclusions

Sprite Sheets limit the Features and Development of the Application

Traditional sprite sheets are limited by the size of spriteSheet the GPU can

handle. This indirectly encourages developers to only support specific resolutions, limit user control over graphics, and to reduce the visual complexity of the application to solve this problem.

7.0 Recommendations

Consider streaming all necessary assets from an external server.

With dynamic sprite sheets, it is a good practice to store every graphic on an external server that can be downloaded by the application only once and cached. These individual files can be allocated to a dynamic spriteSheet only when needed, reducing memory, and improving performancing.

References

- 1 Champion, R., Paci, T. & Vardon, J. (2012). PD 2: Critical Reflection and Report Writing. Retrieved 1 March, 2012 from <https://learn.uwaterloo.ca/d2l/le/content/80224/viewContent/605550/View>
Note: [1] was referenced to format this report.