Diffusertrack: Expressive Manipulation of Diffusion Models in Real-time for Musical Perfromance

DANIEL HEARN

A Thesis Submitted to the Creative Computing Institute
In Particial Fulfillment of the Requirements
for the Degree of Master of Science/MSc

Graduate Programme in Creative Computing

University of the Arts London

ual: creative computing institute

21st October, 2024

Abstract

Example abstract could be: This thesis investigates the expressive use of network manipulation techniques in diffusion models for use in the domain of live musical performance. The study aims to understand how different manipulation techniques can be employed by an artist to generate and control sound in a live setting. Additionally, this study explores how manipulations could be represented in an explainable musical interface. A mixed methods approach was used to gather and analyse data from musicians, artist, and their audiences. The results reveal [key findings]. These findings suggest [implications], providing valuable insights into [broader context]. The study concludes by highlighting the significance of [conclusions] and proposing directions for future research in [related areas].

Acronyms

VR	Virtual Reality
AR	Augmented Reality
CV	Computer Vision
CG	Computer Graphics

Contents

List of Figures

List of Tables

Acknowledgements

Thank you to my supervisors, Prof. Mick Grierson and Dr. Phoenix Perry, along with other members of the CCI, including Dr. Terence Broad, Shuoyang Zheng, and Ilya Pavlov. To my partner Lois and my familiy and friends, thanks for your unconditional support and encouragement.

And finally, to the open-source community, thank you for providing the tools and resources that made this project possible. Together we stand on the shoulders of giants.

Declaration

This thesis has been created by myself and has not been submitted in any previous application for any degree. The work in this thesis has been undertaken by myself except where otherwise stated.

1 Introduction

The introduction provides a concise overview of the research topic, its significance, and relevance. It briefly outlines the research objectives, questions, and hypotheses, and summarizes the structure of the thesis.

1.1 Motivation

In computer animation, when modelling a virtual character, hairstyling is very vital as an essential part for showing a character's personality. As computer graphics hardware and software dramatically improved in today's world, virtual hair modelling is wildly used in film production, video games and other various CG applications. Recently released animation films such as *Frozen* and *Moana* (see Figure 1) benefit from delicate virtual hairstyles to demonstrate the characters.

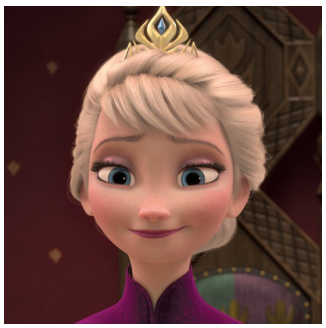

Figure 1: *Frozen hair*

In video games [1], to solve the problem of realism [2], two world leading companies of computer graphics cards - Nvidia & AMD - have both developed their own hair technologies, HairWorks & TressFX respectively (see Figure **??**). They make great efforts to avoid the plastic-molded like hairs appearing in video games. Furthermore, in interactive media and video games, perhaps the most popular facet of any character-customization tool is the ability to customize the appearance of a character in terms of hairstyling, dress and other forms of body modification [3]. All of these reasons make hair modelling recognized as a widespread topic in the computer graphics community.

2 Methodology

The methodology section describes the research design and methods used for data collection and analysis. It includes a brief explanation of the rationale behind the chosen methods, the study's sample, and any instruments or tools used. Ethical considerations and study limitations are also addressed succinctly.

2.1 Sobel Operator

The Sobel operator (Table 1) is represented as two convolving masks along y direction and along x direction for differential operation along y and x direction respectively.

-1	-2	-1
0	0	0
1	2	1

Table 1: *Sobel 2D mask of size 3*3 along y direction*

3 Results

The results section presents the findings of the research clearly and concisely, using tables, graphs, and charts where necessary. Descriptive text accompanies the visual data, highlighting key findings without interpretation.

4 Discussion

In the discussion section, the results are interpreted and analyzed in relation to the research questions and existing literature. This section explores the implications of the findings, addresses any unexpected outcomes, and discusses the study's limitations and strengths.

5 Conclusions

The conclusion section summarizes the main findings and their significance. It restates the research objectives and discusses how they were met, providing a brief

summary of the study's contributions to the field and any practical implications.

6 Future Work

The future work section suggests directions for future research based on the findings and limitations of the current study. It identifies potential areas of improvement and new research questions that have emerged.

References

[1] C. Dobrian and D. Koppelman, "The 'e' in nime: Musical expression with new computer interfaces," Available at: https://escholarship.org/uc/item/8v0022bd [Accessed 21 October 2024], 2006, uC Irvine: Integrated Composition, Improvisation and Technology (ICIT). [Online]. Available: https://escholarship.org/uc/item/8v0022bd

[2] T. Broad, F. F. Leymarie, and M. Grierson, "Network bending: Expressive manipulation of deep generative models," in *Artificial Intelligence in Music, Sound, Art and Design: 10th International Conference, EvoMUSART 2021, Held as Part of EvoStar 2021, Virtual Event, April 7–9, 2021, Proceedings 10*. Springer, 2021, pp. 20–36.

[3] N. Tokui and D. D'Heilly, *Surfing Human Creativity with AI - A User's Guide*. Lulu.com, 2023. [Online]. Available: https://books.google.co.uk/books?id=iNSP0AEACAAJ

Appendices

The appendix includes supplementary material that supports the thesis but is not essential to the main text. This may consist of raw data, detailed methodologies, questionnaires, consent forms, and additional figures or tables.