2D HUMAN ANIMATION SYNTHESIS FROM VIDEOS USING GENERATIVE ADVERSARIAL NEURAL NETWORKS.

Pasan Nadeera Udawatta

199489J

Degree of Master of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa Sri Lanka

March 2022

2D HUMAN ANIMATION SYNTHESIS FROM VIDEOS USING GENERATIVE ADVERSARIAL NEURAL NETWORKS.

Pasan Nadeera Udawatta

199489J

Thesis submitted in partial fulfillment of the requirements for the Degree of Masters of Science in Artificial Intelligence

Degree of Master of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa Sri Lanka

March 2022

Declaration

I declare that this dissertation does not incorporate, without acknowledgment, any material previously submitted for a Degree or a Diploma in any University and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be made available to outside organization.

Name of the Student

P. N Udawatta

Signature of Student

Date:

Supervised by

Dr.Subha Fernando

Signature of the Supervisor

Date:

Dedication

I dedicate this thesis to my parents and my wife who were there for me for my successes and failures.

Acknowledgment

The success of this thesis would not have been possible without the guidance and support of many people around me.

Many thanks to my adviser Dr. Subha Fernando for her effort, patience and wisdom, and guidance provided throughout the project. Her expertise was the guiding light that enabled me to successfully formulate the research question and the methodology.

I would also like to thank Prof. Asoka Karunananda for the valuable guidance that directed me towards the correct path of conducting research.

I would also like to gratefully recognize the efforts of Dr.Sagara Sumathipala for his valuable support to conduct this research promptly. I would also like to offer my gratitude to all my lecturers who provide me with support, guidance, and knowledge that enable me to complete this thesis. I would also like to offer my sincere gratitude to non-academic staff members who were there to help throughout my postgraduate process.

I would also like to recognize my dear colleagues for all the support and guidance provided.

Most importantly, I would like to thank my parents and my wife for their unconditional, unequivocal, and loving support.

Abstract

Synthesizing 2D human animation has many industrial applications yet is currently done manually by animators utilizing time and resources. Therefore, many types of research have been conducted to synthesize human animation using artificial intelligence techniques. However, these approaches lack the quality as well as capability to generalize to various visual styles. Thus, synthesizing high-quality human animations across different visual styles remains a research challenge

We hypothesize that given video references for motion and appearance, synthesizing high-quality human animations across a variety of visual styles can be achieved via generative adversarial networks.

Here we have come up with the solution known as HumAS-GAN, an acronym for **Hum**an **A**nimation **S**ynthesis **G**enerative **A**dversarial **N**etworks. HumAS-GAN accepts video references for motion and appearance and synthesis 2d Human animations. HumAS-GAN has three main modules, motion extraction, motion synthesis, and appearance synthesis. In motion extraction, the motion information is extracted via pre-trained human pose extraction [21], The motion synthesis module syntheses a motion representation matching the target human's body structure which is then combined with the human pose coordinates to be used by the appearance synthesis module to generate the Human animation. HumAS-GAN is focused on improving the quality of the animation as well as the ability to use cross-domain/visual-style references to generate animation. This solution will be beneficial for many multimedia-based industries as it is capable of generating high human animations and quickly switching to any visual style they prefer.

HumAS-GAN is evaluated against other methods using a custom dataset and a set of 3 experiments designed to evaluate the capability of generating human animations across various visual styles. Evaluations results prove the superiority of HumAS-GAN over other methods in synthesizing high-quality 2d human animations across a variety of visual styles.

Contents

List of tables

List of figures

Abbreviations

GAN	Generative Adversarial Network
CGI	Computer Generated imagery
HumAS-GAN	Human Animation Synthesis Generative Adversarial Network
cGAN	Conditional Generative Adversarial Network
GPU	Graphics Processing Unit
HD	High Definition
SSIM	Structural Similarity Index Measure
LPIPS	Learned Perceptual Image Patch Similarity