CAIRO UNIVERSITY

FACULTY OF ENGINEERING

DEPARTMENT OF COMPUTER ENGINEERING

Retratista

A Graduation Project Report Submitted

to

Faculty of Engineering, Cairo University

in Partial Fulfillment of the requirements of the degree

of

Bachelor of Science in Computer Engineering.

Presented by

Mohamed Shawky Zaky AbdelAal Sabae

Remonda Talaat Eskarous

Mohamed Ahmed Mohamed Ahmed

Mohamed Ramzy Helmy Ibrahim

Supervised by

Dr. Mayada Hadhoud

26th July, 2021

All rights reserved. This report may not be reproduced in whole or in part, by photocopying or other means, without the permission of the authors/department.

Abstract

الملخص

Acknowledgement

Contents

Abstract (English) **1**

Abstract (Arabic) **2**

Acknowledgement **3**

Table of Contents **5**

List of Figures **6**

List of Tables **7**

List of Abbreviation **8**

List of Symbols **9**

Contacts **10**

1 Introduction **11**

2 Market Feasibility Study **12**

3 Literature Survey **13**

4 System Design and Architecture **14**
4.1 Overview and Assumptions 14
4.2 System Architecture 16
4.2.1 Block Diagram 16
4.3 Module 1 : Speech Recognition 18
4.3.1 Functional Description 18
4.3.2 Modular Decomposition 18
4.3.3 Design Constraints 18
4.3.4 Other Description 18
4.4 Module 2 : Text Processing 18
4.4.1 Functional Description 18
4.4.2 Modular Decomposition 18
4.4.3 Design Constraints 18
4.4.4 Other Description 18
4.5 Module 3 : Face Code Generation 18
4.5.1 Functional Description 18
4.5.2 Modular Decomposition 18
4.5.3 Design Constraints 18
4.5.4 Other Description 18
4.6 Module 4 : Code-to-Face Translation 18
4.6.1 Functional Description 18

4.6.2 Modular Decomposition 18
4.6.3 Design Constraints 18
4.6.4 Other Description 18
4.7 Module 5 : Face Refinement 18
4.7.1 Functional Description 18
4.7.2 Modular Decomposition 18
4.7.3 Design Constraints 18
4.7.4 Other Description 18
4.8 Module 6 : Multiple Head Poses Generation 18
4.8.1 Functional Description 18
4.8.2 Modular Decomposition 18
4.8.3 Design Constraints 18
4.8.4 Other Description 18

5 System Testing and Verification 19

6 Conclusions and Future Work 20

A Development Platforms and Tools 21

B Use Cases 21

C User Guide 21

D Code Documentation 21

E Feasibility Study 21

List of Figures

4.1 Block diagram of complete system architecture 16
4.2 Block diagram of application design 16

List of Tables

List of Abbreviation

List of Symbols

Contacts

1 Introduction

2 Market Feasibility Study

3 Literature Survey

4 System Design and Architecture

In this chapter, we discuss our working pipeline and system architecture in details. Generally, our system takes a speech note, textual description or numerical attributes as an input. It processes the input description and outputs the initial human face portrait that corresponds to the given description. Afterwards, the user is allowed to manually control some facial attributes and morphological features and to rotate the face and render it in multiple poses. In the first section, we give an overview about the system. Then, we discuss the system architecture in the second section. In the subsequent sections, each module implementation is discussed in details.

4.1 Overview and Assumptions

As mentioned above, our system basically enables the user to describe a human face in words or using numerical values and turns it into a full human face portrait that can be manipulated and rendered in multiple poses. The system relies heavily on generative models and text processing, both are iteratively designed to obtain the required results. The overall flow can be described as follows :

- The input speech notes are translated to text.
- The textual description (extracted from speech input or manually entered) is processed to extract the numerical values of the required facial features.
- The numerical values are used generate a face embedding vector that encodes the facial attributes in low dimensional space ($512D$).
- A generative model is specifically designed to translate from the low dimensional embedding into the full face portrait ($1024X1024$).
- The generated face portrait can be further refined by navigating the face embedding space and re-generating the face portrait.
- Once the user settles on the final face portrait, the system can render that face in multiple poses to provide further identification.

The previous flow provides a very versatile framework to generate face portrait and adjust it to your liking. However, there is an extremely large number of facial attributes and morphological features to describe a human face. Consequently, we have to choose a descriptive subset of these attributes to consider in the face description. We consider 32 facial attributes for face description, which are listed as follows :

- Overall face :
  - Gender : Male / Female.
  - Age : Young / Old.
  - Thickness : Chubby / Slim.

  - Shape : Oval / Circular.
  - Skin Color : Black / White.
  - Cheeks : Normal / Rosy.
- Eyes :
  - Color : Black / Blue / Green / Brown.
  - Width : Wide / Narrow.
  - Eyebrows : Light / Bushy.
  - Bags Under Eyes : On / Off.
- Nose :
  - Size : Big / Small.
  - Pointy : On / Off.
- Ears :
  - Size : Big / Small.
- Jaw :
  - Mouth Size : Big / Small.
  - Lips Size : Big / Small.
  - Cheekbones : Low / High.
  - Double Chin : On / Off.
- Hair :
  - Color : Black / Blonde / Brown / Red / Gray.
  - Length : Tall / Short.
  - Style : Straight / Curly / Receding Hairline / Bald / with Bangs.
- Facial Hair :
  - Beard / None.
- Race :
  - White / Black / Asian.
- Accessories :
  - Glasses : Sight / Sun.
  - Makeup : On / Off.
  - Lipstick : On / Off.

4.2 System Architecture

Now, let's discuss our system architecture. The system consists of 7 modules, 3 core modules of the project and 4 auxiliary modules.

4.2.1 Block Diagram

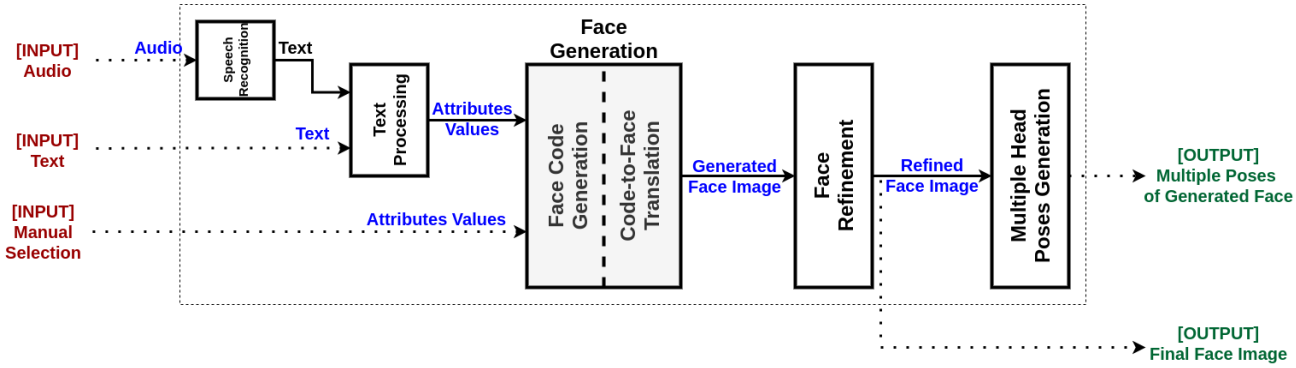

Figure 4.1: Block diagram of complete system architecture

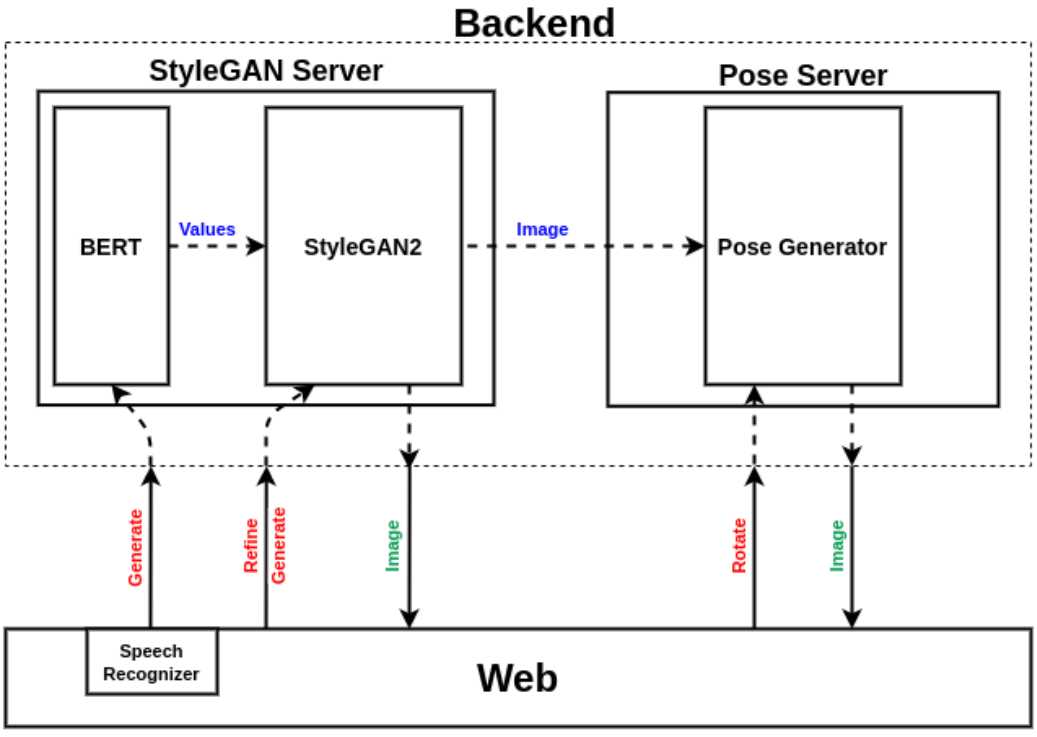

Figure 4.2: Block diagram of application design

4.3 Module 1 : Speech Recognition

4.3.1 Functional Description

4.3.2 Modular Decomposition

4.3.3 Design Constraints

4.3.4 Other Description

4.4 Module 2 : Text Processing

4.4.1 Functional Description

4.4.2 Modular Decomposition

4.4.3 Design Constraints

4.4.4 Other Description

4.5 Module 3 : Face Code Generation

4.5.1 Functional Description

4.5.2 Modular Decomposition

4.5.3 Design Constraints

4.5.4 Other Description

4.6 Module 4 : Code-to-Face Translation

4.6.1 Functional Description

4.6.2 Modular Decomposition

4.6.3 Design Constraints

4.6.4 Other Description

4.7 Module 5 : Face Refinement

4.7.1 Functional Description

4.7.2 Modular Decomposition

4.7.3 Design Constraints

4.7.4 Other Description

4.8 Module 6 : Multiple Head Poses Generation

4.8.1 Functional Description

4.8.2 Modular Decomposition

4.8.3 Design Constraints

4.8.4 Other Description

5 System Testing and Verification

6 Conclusions and Future Work

A Development Platforms and Tools

B Use Cases

C User Guide

D Code Documentation

E Feasibility Study