Faculty of Mathematics and Information Science Warsaw University of Technology

Biometric fusion system for human recognition using face and voice

prepared by:
Kuśmierczyk Aleksander
Sławińska Martyna
Żaba Kornel

Version no. 1.1

Date: 15.11.2017

History of changes

Example:

| Date	Author	Description	Version
14.10.2017	Kuśmierczyk Aleksander Sławińska Martyna Żaba Kornel	First version	1.0
19.10.2017	Kuśmierczyk Aleksander Sławińska Martyna Żaba Kornel	Glossary added	1.0.1
15.11.2017	Kuśmierczyk Aleksander Sławińska Martyna Żaba Kornel	Technical project	1.1

Table of Contents

1 Specification

1.1 Executive summary

This system is designed for identification and/or verification of human. With the use of biometric algorithms based on extracting face and voice features, the program recognizes the user with the sufficient degree of compliance. The acquisition of the aforementioned data is done by the portable or built-in camera with a microphone.

1.2 Functional requirements

1. The system should display graphical user interface with options.

2. The system should allow the user to choose what to identify/verify first: face or voice.

3. The system should identify and/or verify the face:

    3.1. The system should use the camera for capturing the image of the face.

    3.2. The system should compare PV's biometric with the biometrics stored in the database.

4. The system should identify and/or verify the speech/voice:

    4.1. The system should use the microphone for recording the speech.

    4.2. The user should say the specific word in order for the system to identify his/her voice.

    4.3. The system should compare PV's biometric with the biometrics stored in the database.

5. After collecting PV's face and speech data, the system should compare collected results with the database and give an answer whether the PV already belongs to the database collection (verifying stage).
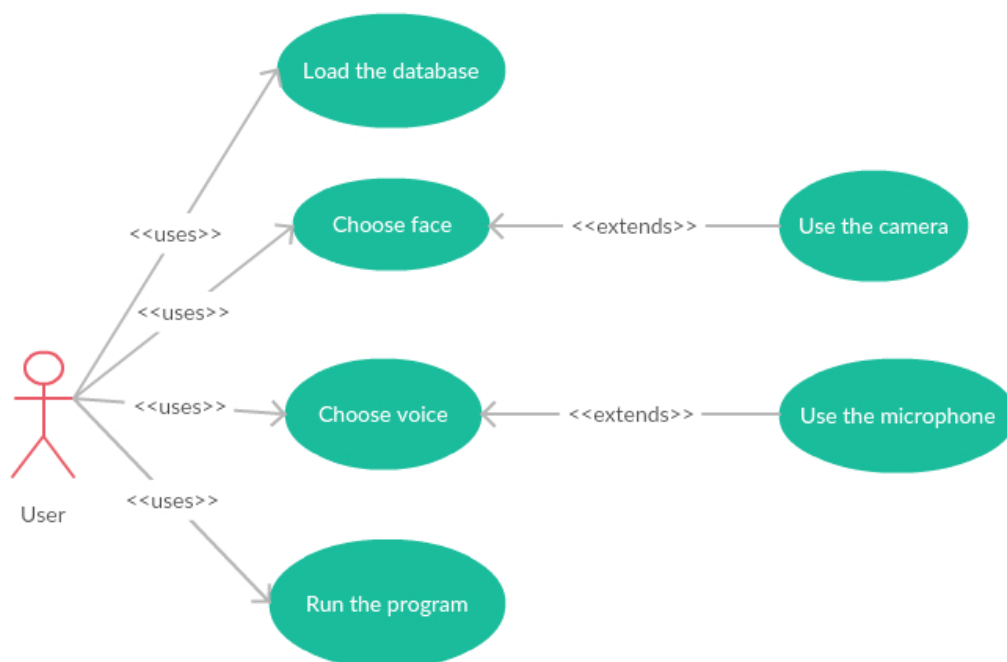
Actor	Name	Description	System response
User	Load the database	Load the database for machine learning purposes	The system loads the database from location pointed by the user.
	Choose face	Choose face for identification	The program starts the camera
	Choose voice	Choose voice for identification	The program starts the microphone.
	Run the program	Run the program which compares the PV's data with the database collection	The system returns the result of recognition.
	Use the camera	Use the camera for capturing the face of PV	The system captures the face of PV with camera.
	Use the microphone	Use the microphone for recording the speech of PV	The system records the voice of PV with microphone.

User stories

1. Attempt of PV verification

1.1   The user starts the program, chooses which biometric should be used in recognition process as first. Then the user uses the appropriate hardware to capture the required input or uses the previously prepared data for the program to use in order to recognize the PV. The user reads the results of the verification.

1.3 Non-functional requirements
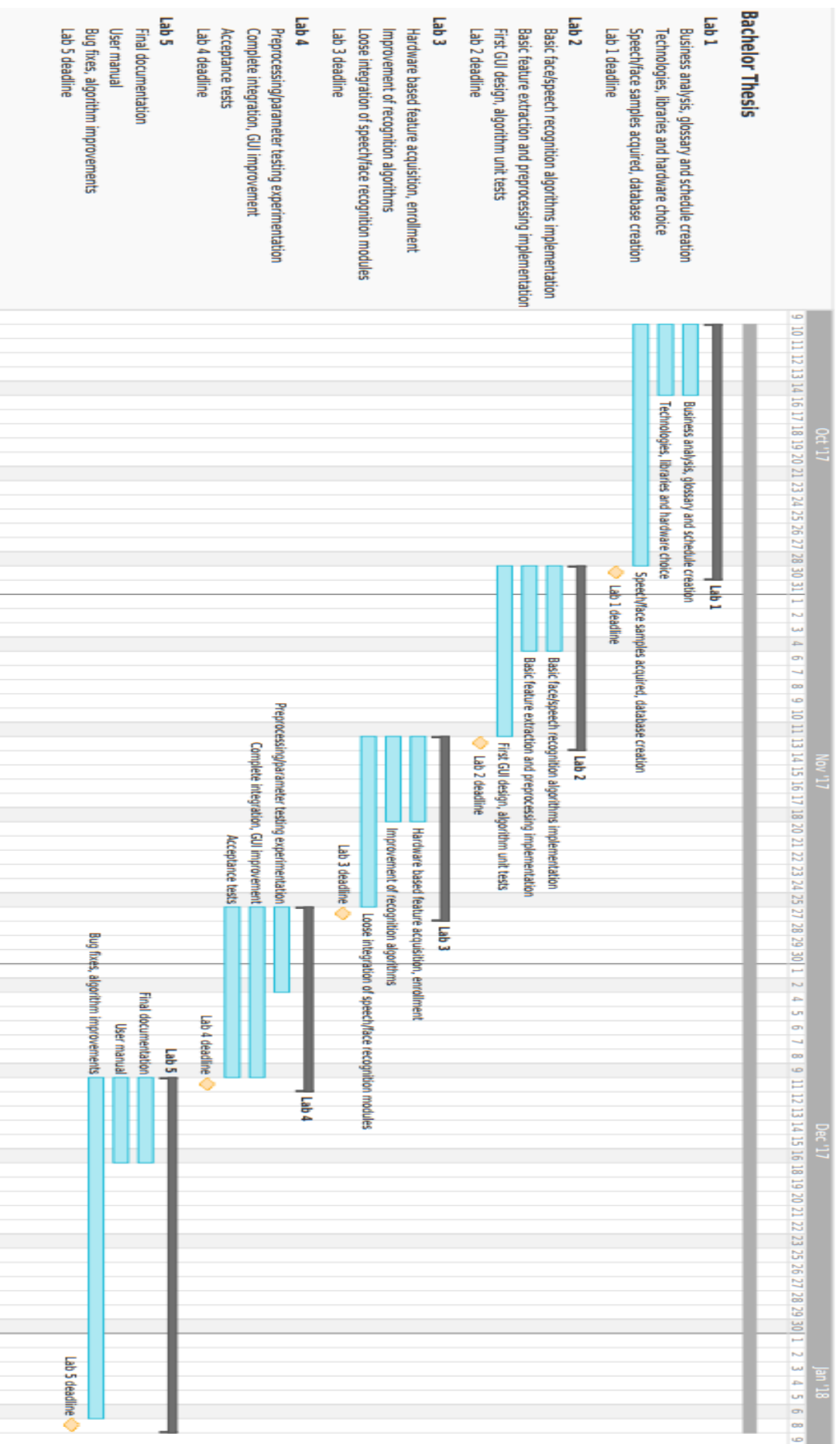
1. The system does not have any login options.

2. The system allows identifying one face/voice at the time.

3. The system should support the Windows operating system.

4. The system should work in an offline mode.

5. The system response time depends on the size of the database.

6. The system's hardware connects by USB cable: camera and microphone, or is built-in the workstation.

Below is the table that describes the URPS requirements of the program.

Requirements	Nr wymagania	Opis
Usability	1	The software is user friendly, as the only requirement for the use of the program is to load the database and to capture the image or to record the speech.
Reliability	2	The system is a desktop application, after manual start-up it is ready to work.
	3	The program handles the incorrect data types in the database, and shuts down if a hardware malfunction occurs.
Performance	4	The system returns the verification result with the delay basing on the database size.
	5	The program verifies the PV using two biometric features, thus achieves the sufficiently good verification correctness.
Supportability	6	In order to use the application, the user is required to have Windows operating system.
	7	The program requires a built-in or a USB-connected camera with a microphone.
	8	In order for the program to function correctly, it requires a database consisting of a large enough samples of face images and voice recordings.

1.4 Project schedule

DEADLINE	MILESTONE	LAB DATE
2017-10-14	Business analysis, glossary and schedule creation	LAB 1 30.10.2017: Business analysis, Glossary, Supplementary spec, Schedule
2017-10-14	Technologies, libraries and hardware choice	
2017-10-28	Speech/face samples acquired, database creation	
2017-11-04	Basic face/speech recognition algorithms implementation	LAB 2 13.11.2017: Technical project, Choice of development model
2017-11-04	Basic feature extraction and pre-processing implementation	
2017-11-11	First GUI design, algorithm unit tests	
2017-11-18	Hardware based feature acquisition, enrolment	LAB 3 27.11.2017: Source code of non-integrated modules, unit tests
2017-11-18	Improvement of recognition algorithms	
2017-11-25	Loose Integration of speech/face recognition modules.	
2017-12-02	Pre-processing/parameter testing experimentation	LAB 4 11.12.2017: Integrated modules, smoke test, acceptance tests
2017-12-09	Complete integration, GUI improvement	
2017-12-09	Acceptance tests	
2017-12-16	Final documentation	LAB 8.01.2018: Working v.1.0, user manual, code/docs, acceptance tests results, list of changes
2017-12-16	User manual	
2018-01-06	Bug fixes, algorithm improvements	
2018-01-06	Others	

Bachelor Thesis

Lab 1
Business analysis, glossary and schedule creation
Technologies, libraries and hardware choice
Speech/face samples acquired, database creation
Lab 1 deadline

Lab 2
Basic face/speech recognition algorithms implementation
Basic feature extraction and preprocessing implementation
First GUI design, algorithm unit tests
Lab 2 deadline

Lab 3
Hardware based feature acquisition, enrollment
Improvement of recognition algorithms
Loose integration of speech/face recognition modules
Lab 3 deadline

Lab 4
Preprocessing/parameter testing experimentation
Complete integration, GUI improvement
Acceptance tests
Lab 4 deadline

Lab 5
Final documentation
User manual
Bug fixes, algorithm improvements
Lab 5 deadline

1.5 Glossary

Biometric – feature owned by human, which allows verification of an individual

Enrolment – process of acquisition of data of the new PV

Face features – face biometric

Feature acquisition – process of collecting data

Feature extraction – process of extracting biometric information from data

Identification – the automatic identification of living individuals by using their physiological and behavioural characteristics

Offline mode – the mode of program, which does not require internet connection

PV – person being verified

Recognition - refers to the automated recognition of individuals based on their biological and behavioural traits

Verification - an identity authentication process used to confirm a claimed identity through uniquely identifiable biological traits

Voice features – voice biometric

2 Technical project

2.1 Selected technology

The technology we have chosen for our project is the **C# .NET 4.6** due to the familiarity with this framework of our group. Furthermore, the availability of the online sources of information for this technology allowed us to search for the solutions of encountered implementation problems in the web. Throughout the work it has proven to be stable and provided useful functionalities.

For the development and management of the database used in this project we have chosen **MSSQL Server 2012**. This platform has proven to be stable and using the 2012 version allowed us to find multiple solutions to the problems we have encountered during the work over the project. Besides, our group is familiar to it's structure due to the previous experience with this program.

For capturing of images and the recording of sound we have chosen to use external program - **FFmpeg** as it provided us with easy to adapt commands for the aforementioned activities. As hardware solution we have chosen Logitech HD Webcam C270 as it provides the necessary video and audio data required by the project.

2.2 Development model

The goal of our project is to provide the method of identification and/or verification of a person, the study over this subject is still being developed, thus the model we had to adapt for its creation had to be flexible and allow us for frequent theoretical discussion. The development technique we have chosen is the **iterative waterfall model** with the elements of extreme programming. Such choice is based on the need of flexibility for implementation and theoretical discussion, which this model provides, as well as the constant assessment of the chosen methods and their adaptation. All the methods implemented are tested afterwards and the results are discussed with the supervisor in order to gain more understanding over them and to establish further direction for the project. Elements of extreme programming we have adapted to our model were the periodic replacement of the roles of people in the team. Such decision is made in order to allow all the members to follow and to understand the theoretical reasons behind the decisions over the methods of human recognition. Furthermore, this supports better quality of the code produced and allowed for frequent discussion over the problems encountered during the development process.

2.3 System architecture

System architecture is composed of layered modules, each representing different responsibility.

List of modules:

Common - module responsible for database access, as well as containing utility classes used by multiple of below modules.

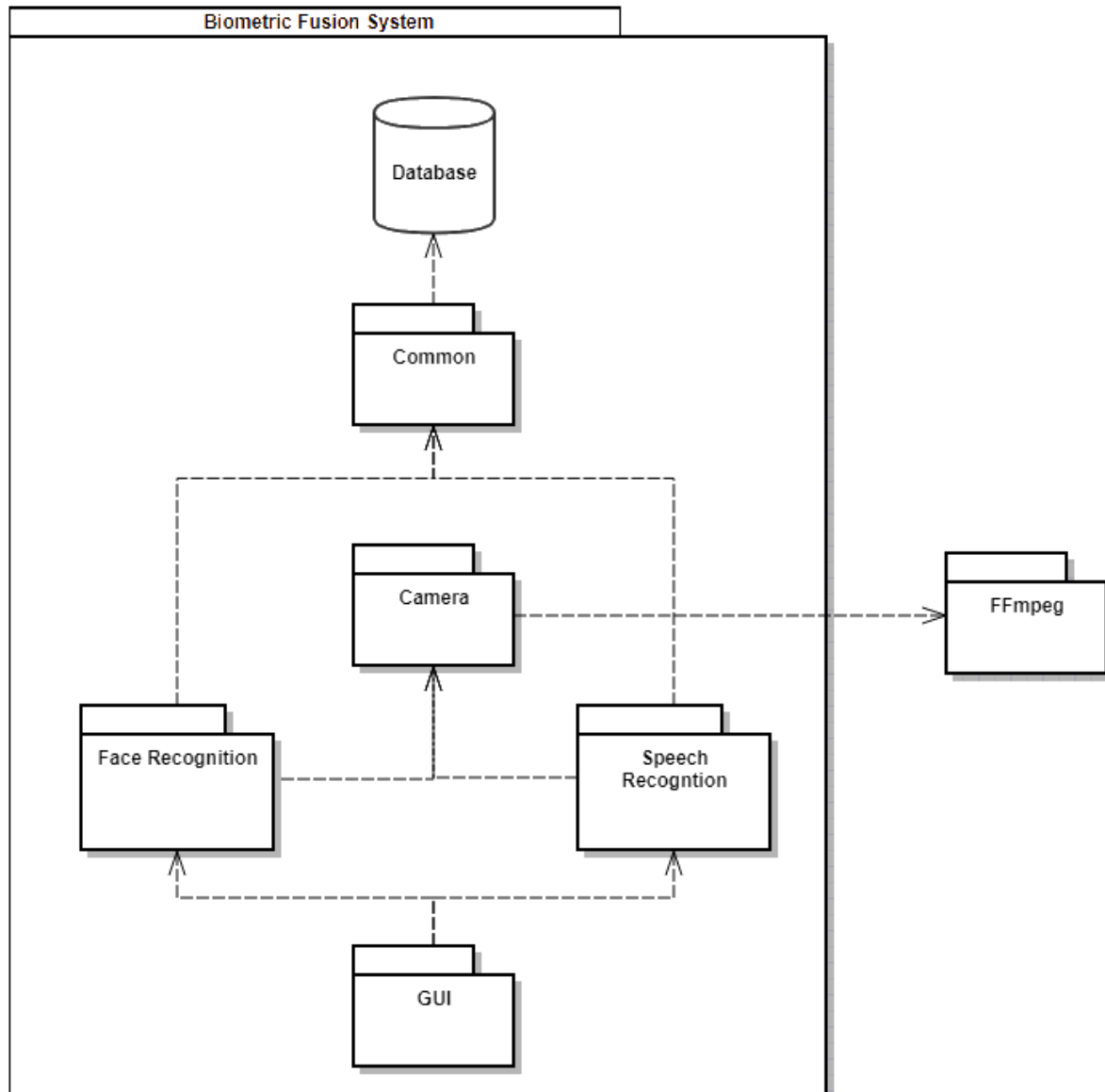**Camera** - module responsible for acquisition of the face images and speech recordings. Depends on the scriptable, external video and audio recording program **FFmpeg**.

FaceRecognition - module containing algorithms for face image preprocessing, feature extraction and recognition.

SpeechRecognition - module containing algorithms for speech recording preprocessing, feature extraction and recognition.

Gui - module defining user interface and interaction. Depends on the .NET desktop application framework **Windows Presentation Foundation.**

Package diagram of system modules:

2.4 Class diagram

The below picture represents the class diagram of the whole system.

Camera

- picture : Bitmap

+ TakePicture (picture) : void
+ GetPicture ( ) : Bitmap

Microphone

- speech : Audio

+ RecordSpeech (speech) : void
+ GetSpeech ( ) : Audio

<uses>

<uses>

ImageLoader

- picture : Bitmap

+ Load (picture) : void

AudioLoader

- speech : Audio

+ Load (speech) : void

<uses>

<uses>

FaceFeatureExtraction

- picture : Bitmap
- features : List<Vector>

+ Extract (picture, features) : void

SpeechFeatureExtraction

- speech : Audio
- features : List<Vector>

+ Extract (speech, features) : void

FaceComparison

- dbFaces : List<Feature>
- faceToCompare : Feature

+ Compare (dbFaces, faceToCompare) : bool

SpeechComparison

- dbSpeeches : List<Feature>
- speechToCompare : Feature

+ Compare (dbSpeeches, speechToCompare) : bool

Feature

- f : int []

+ GetFeature ( ) : Vector

<uses>

<uses>

PersonVerificator

+ IsVerified ( ) : bool