Plataforma web para el análisis de la gesticulación facial

Jaime Jesús Delgado Meraz
Advisor: Pedro Luis Sánchez Orellana
Co-advisor: José Torres Jiménez

Maestría Profesionalizante en Sistemas Computacionales
División de Estudios de Posgrado de Investigación
Instituto Tecnológico de Ciudad Victoria

June 22, 2012

Table of Contents

What is Face Analysis?

- The analysis of facial expressions constitutes a critical and complex portion of our non-verbal social interactions.

Facial Expression Recognition is focused on classification of the structure of the face into a set of general emotions.

Facial Expresion Analysis is focused on meassuring how these emotions are produced in the face.


sadness
1 drooping upper eyelids
2 losing focus in eyes
3 slight pulling down of lip corners
anger
1 eyebrows down and together
2 eyes glare
3 narrowing of the lips
contempt
1 lip corner tightened and raised on only one side of face
disgust
1 nose wrinkling
2 upper lip raised
surprise
Lasts for only one second:
1 eyebrows raised
2 eyes widened
3 mouth open
fear
1 eyebrows raised and pulled together
2 raised upper eyelids
3 tensed lower eyelids
4 lips slightly stretched horizontally back to ears

Facial Gestures Recognition

- There are three main steps for working in facial gestures recognition:
  1. Face detection,
  2. Extraction of facial gestures information, and
  3. Classification of the information extracted.

  .
- The methods for facial gesture analysis can be grouped in **deformation** extraction and **motion** extraction.

Motivation

- Human interaction is very important in many different aspects in our daily life, such as business meeting, medical diagnostics, etc.
- However, thanks to the improvement of technology, this operations can be done remotely, which can be especially useful when the distance is considered a barrier, but this producess another problem which is the lack of physical gesture apreciation.

Figure: Example of the lack of physical gesture apreciation

Problem Statement

Problem

Given a sequence of images from the face, taken in a client computer, where the illumination conditions and background motion are restricted, the problem of analyzing facial gestures on the web, could be stated as an architecture that allows transmission of compressed images through a web service, in order to produce an analysis of facial gestures.

- From this we state the following research questions:
  1. Considering the speed of internet connections in México, what techniques can be used to increase the speed of image transfer?
  2. Which type of web service is more suitable for image transmission/analysis?

Objectives

General

Design a web architecture that allows the efficient transmission of facial image sequences to compute an analysis about the behaviour of facial gesticulations, in order to obtain symmetrical / asymmetrical motion patterns which might be used for the decision-making process.

Objectives (cont.)

Specific

1. Implement an image compression technique that allows fast Internet broadcasting.
2. Design and implement a web architecture for the reception and decompression of compressed images.
3. Adapt a given algorithm for the analysis of gesticulations in the transmitted images.
4. Design a set of experiments to test the web architecture.

Scope and Constraints

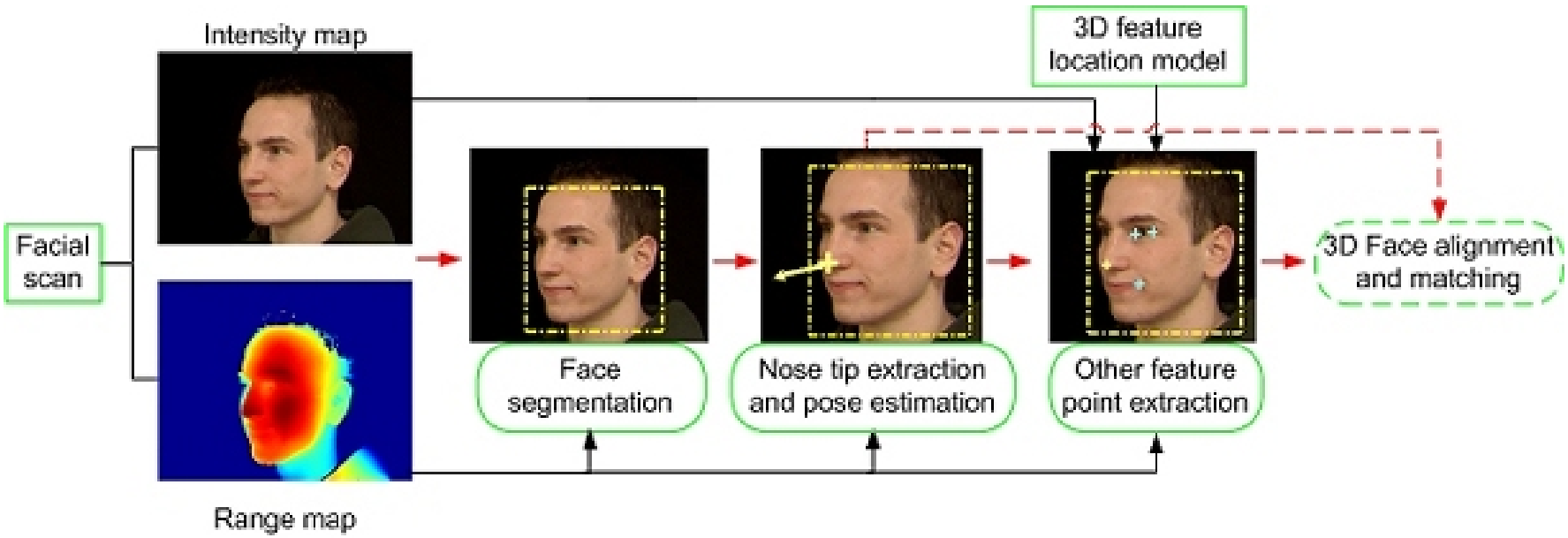
- The proposed architecture will focus only on the compression-transmission process of images, it will not consider the motion from other regions of the body.
- Moreover, the mechanisms for the analysis of the image sequences were generated in collaboration with researchers of the Universidad Politécnica de Victoria and will only be used, but not implemented.

Table of Contents

Facial Gesture Analysis Techniques

- The analysis problem has been approached from two main streams [3]: facial deformation extraction models and facial motion extraction models (both 2D or 3D).
- **2D based techniques**.
  Advantages: fast processing. Disadvantages: very large training data sets is needed to model environment variations.
- **3D based approaches**.
  Advantages: higher accuracy [2]. Disadvantages: Higher computational complexity.

- Therefore...


Intensity map
3D feature location model
Facial scan
Face segmentation
Nose tip extraction and pose estimation
Other feature point extraction
3D Face alignment and matching
Range map

Spiking Neural Networks

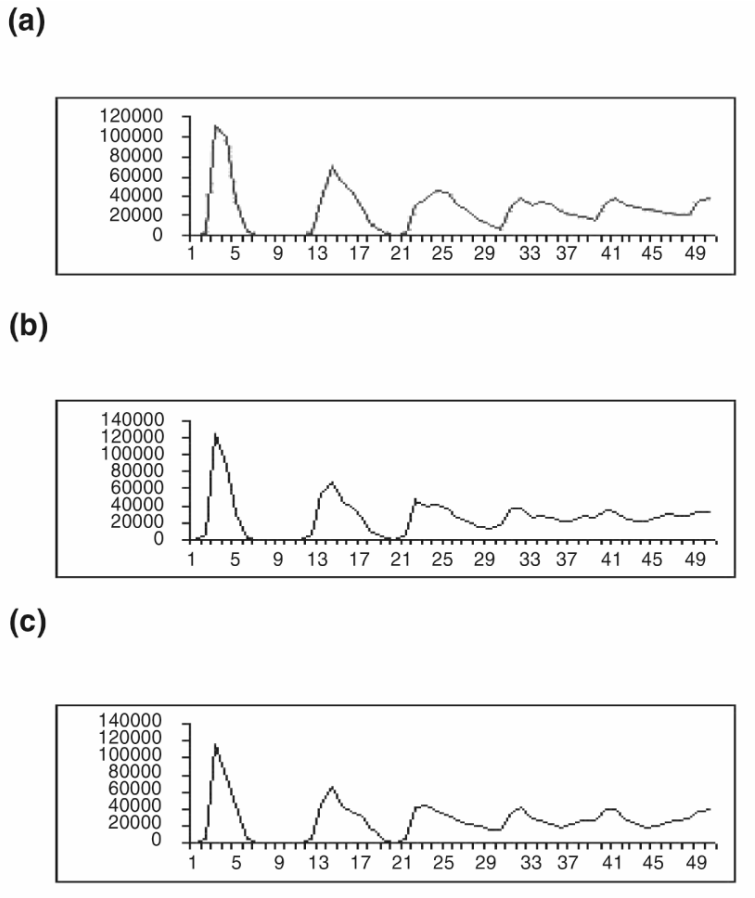


Figure: Obtained results: (a) Left-Right Eyes, (b) Left-Right Nose, (c)Left-Right Mouth

Web Services

- SOAP: Simple Object Access Protocol [8]. It was designed to be a platform and language-neutral alternative to previous middleware technologies.
- REST: Representative State Transfer [4]. A later evolution to SOAP disadvantages.

- Since 3D images are heavier, it is necessary to implement an image compression technique. Basically they can be divided into:
- Lossless Compression Tecniques: The feature of the lossless compression technique is that the original image can be perfectly recovered from the compressed image [5] (such as Run length encoding, Huffman encoding, etc.).
- Lossy Compression Technique provide higher compression ratio than lossless compression (such as Vector quantization, Fractal coding, etc.).

- The experimental design is a key piece in the software development process given that it allows identifying failures in the software before it begins to operate.
- An alternative strategy to test software is the use of Covering Arrays (CAs).
- A Covering Array is a combinatorial object that, with a small number of cases, covers a certain level of interaction of a set of parameters [7].

- A covering array is an $N \times k$ matrix over an alphabet v each $N \times k$ subset contains at least one time each combination from $0, 1, \ldots, v-1^t$, given a positive integer value t.

- Covering Arrays have been an object of study and application in different research areas. Cawse [1] used CAs in the material design, Hedayat et al. [6] used them in medicine and agriculture.

Table of Contents

Methodology

1st Stage In this stage, the different techniques for image compression/decompression will be studied, in order to choose a set of them to be implemented and tested. Once the implementations of the selected techniques have been made, they will be tested and compared to choose only one that will be integrated into a client application. This client application will be responsible of capturing the images for their compression by the selected technique.

Methodology

2nd Stage In this stage, the different approaches for web services will be analized, to choose the most appropriate solution for the problem to be solved. Also, a local web server will be mounted, to implement the chosen web service model...

Methodology

3rd Stage In this stage, the work will be focused on the adaptation of the algorithm for facial gestures analysis, since it was originally developed to work in a local environment. This adaptation will include a mechanism to interact with the compression/decompression technique previously developed.

Finally, a set of experiments oriented to test the different components of the architecture (such as compression, quality of results, processing time, etc.) will be developed and applied.

Table of Contents

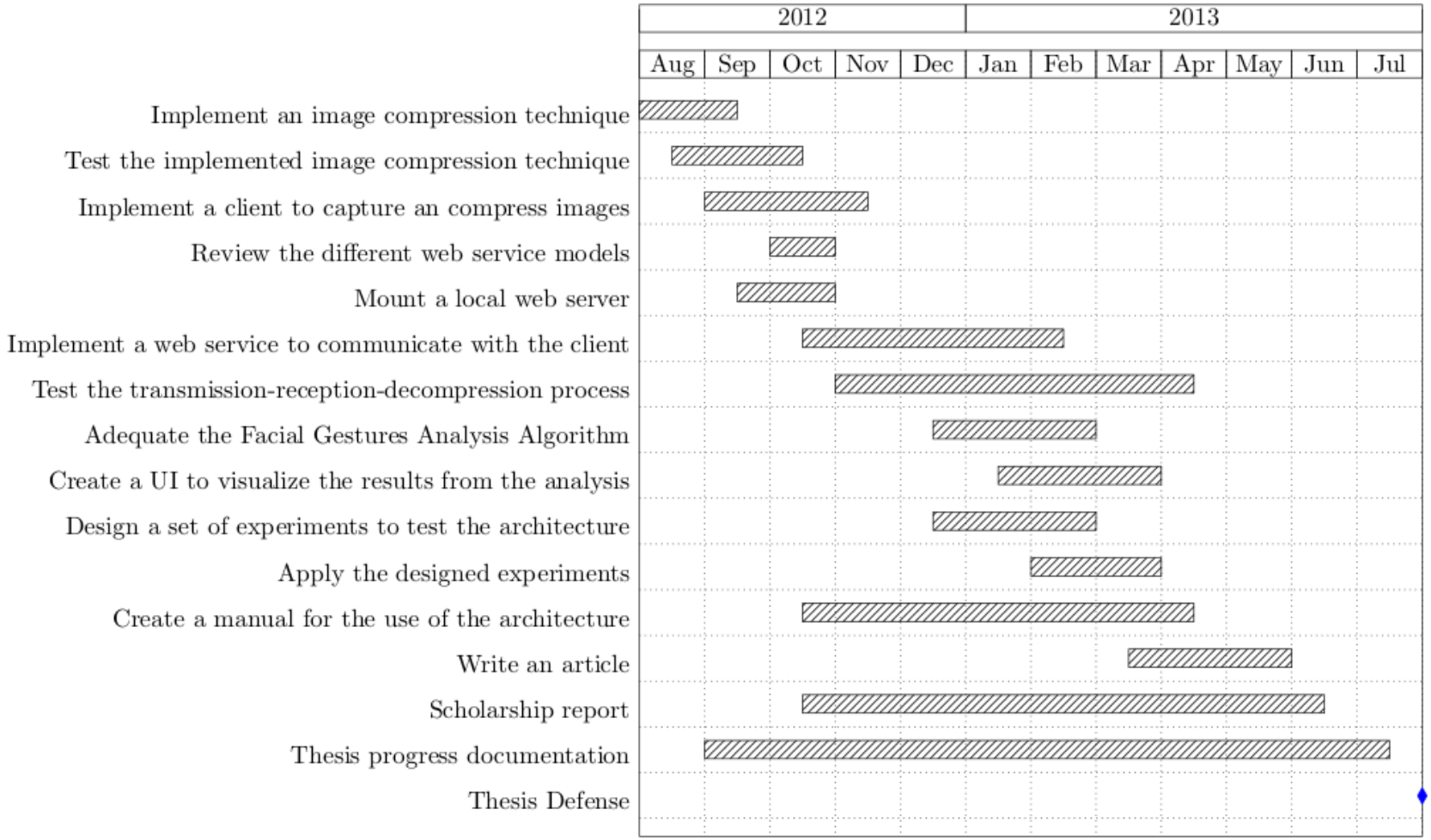

2012
2013
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Implement an image compression technique
Test the implemented image compression technique
Implement a client to capture an compress images
Review the different web service models
Mount a local web server
Implement a web service to communicate with the client
Test the transmission-reception-decompression process
Adequate the Facial Gestures Analysis Algorithm
Create a UI to visualize the results from the analysis
Design a set of experiments to test the architecture
Apply the designed experiments
Create a manual for the use of the architecture
Write an article
Scholarship report
Thesis progress documentation
Thesis Defense

Table of Contents

Infraestructure

- Laptop HP Pavilion DV5-1135la

  OS Ubuntu Linux 11.10 (32 bits).

  Microprocessor AMD Turion X2 Dual-Core Mobile Processor RM-72 (2.1 GHz).

  Memory 3072MB 800MHz DDR2 System Memory (2 Dimm).

  Video Graphics ATI Radeon HD 3450 Graphics (1533MB total graphics memory with 256MB dedicated).

  Hard Drive 320 GB (5400 RPM).

  Webcam Webcam HP Pavilion (VGA low-light)

  Resolution 648×480.

  FPS 30 fps.

- Both the client and the server, will be developed on open source software.

Table of Contents

References I

Cawse, J.
Experimental design for combinatorial and high throughput materials development.
Wiley-Interscience, 2003.

Fang, T., Zhao, X., Ocegueda, O., Shah, S. K., and Kakadiaris, I. A.
3D/4D facial expression analysis: An advanced annotated face model approach.
Image and Vision Computing (2012).

Fasel, B., and Luettin, J.
Automatic facial expression analysis: a survey.
Pattern Recognition 36, 1 (2003), 259–275.

References II

Fielding, R. T.
REST: Architectural Styles and the Design of Network-based Software Architectures.
Doctoral dissertation, University of California, Irvine, 2000.

González, R., and Woods, R.
Digital Image Processing.
Pearson/Prentice Hall, 2008.

Hedayat, A., Sloane, N., and Stufken, J.
Orthogonal Arrays: Theory and Applications.
Springer Series in Statistics. Springer, 1999.

References III

Kuhn, R., Lei, Y., and Kacker, R.
Practical combinatorial testing: Beyond pairwise.
IT Professional 10 (2008), 19–23.

Pautasso, C.
SOAP vs. REST - bringing the web back into web services.
Building (2007), 1–19.