# An approach for Brazilian Sign Language (BSL) recognition based on facial expression and k-NN classifier

Tamires Martins Rezende[1]
Cristiano Leite de Castro[1]
Sílvia Grasiella Moreira Almeida[2]

[1]The Electrical Engineering Graduate Program - UFMG
[2]Department of Industrial Automation - IFMG Ouro Preto.

# Summary

## Brazilian Sign Language (BSL) or Libras

- The second official language of Brazil.

- 5 main parameters:
  1. point of articulation,
  2. hand configuration,
  3. movement,
  4. palm orientation and
  5. non-manual expressions.

- This paper did an exploratory study of the peculiarities involved in non-manual sign language expression recognition.

## Brazilian Sign Language (BSL) or Libras

The BSL recognition using computational methods is a challenge
for a variety of reasons:

- There is currently no standardized database containing signs
  in a format that allows computer classification systems
  validation;

- One sign is composed of simultaneous elements;

- The language does not contain a consistent identifier for the
  start and end of a sign;

- Different people complete any given gesture differently.

# The Brazilian Sign Language Database

The first step was to choose the signs that contained changes in facial expression during execution.
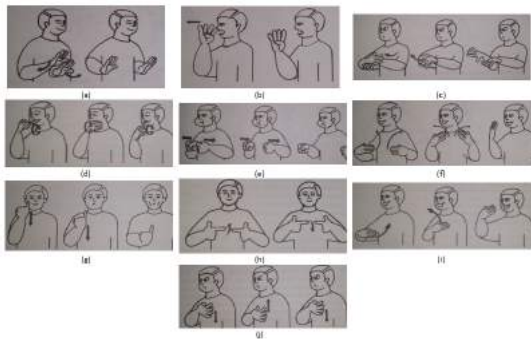


Figure: Signs: (a) to calm down, (b) to accuse, (c) to annihilate, (d) to love, (e) to gain weight, (f) happiness, (g) slim, (h) lucky, (i) surprise, and (j) angry

## Record Protocol

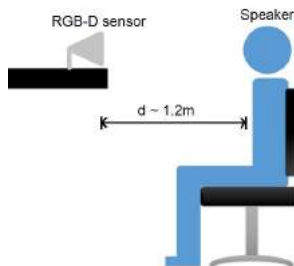After first step to record, the signs were recorded.



Figure: Scenario created for record the signs

With the 10 signs each recorded 10 times with the same speaker, the balanced database had a total of 100 samples.

## Record Protocol

RGB-D sensor $\Rightarrow$ Kinect
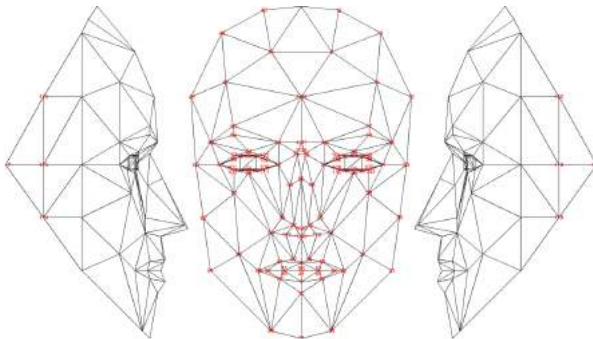Software $\Rightarrow$ nuiCaptureAnalyse



Figure: Facial model used with labeled points
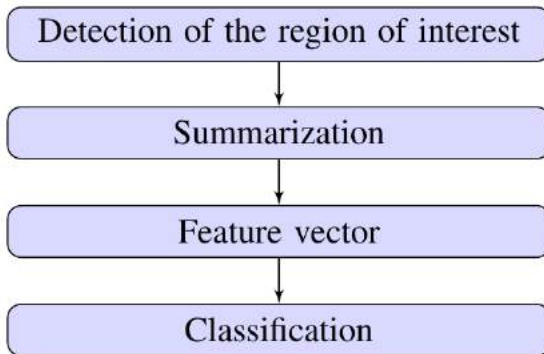
## Methodology



Figure: Methodology for the automatic face recognition system

## Step 1: Detection of the region of interest

**Input**: Signals recorded by Kinect.
**Output**: Video with the face images.



Figure: Full frame



Figure: Selected face

## Step 2: Summarization

**Input**: Face images.
**Output**: 5 most relevant frames.



Figure: The five most relevant frames extracted from a recording of the sign "to love"

## Step 3: Feature vector

**Input**: 5 most relevant frames.
**Output**: Descriptor of a sign.

Descriptor of a single frame:

$$D = \left[ \begin{array}{cccccccc} x_1 & y_1 & x_2 & y_2 & ... & x_{121} & y_{121} \end{array} \right]_{1 \times 242}$$

Sign:

$$Vector = \left[ \begin{array}{ccccc} D_1 & D_2 & D_3 & D_4 & D_5 \end{array} \right]_{1 \times 1210}$$

$$Vector = \left[ \begin{array}{cccccccc} x_{1,1} & y_{1,1} & ... & x_{i,j} & y_{i,j} & ... & x_{5,121} & y_{5,121} \end{array} \right]_{1 \times 1210}$$

$$i = \text{number of the frame and}$$
$$j = \text{point of the face}$$

## Step 4: Classification

**Input**: Descriptor of a sign.
**Output**: Average accuracy of the 10 iterations.

**Algorithm 1: K-NN CLASSIFICATION**

Input: Sign samples
Output: $acc_{avg}$ and $\sigma$ of the 10 iterations

1 **Start**
2    **for** $w = 1$ to $10$ **do**
3      Randomizes the samples of each sign
4      $train \leftarrow 80\%$ of the data
5      $test \leftarrow 20\%$ of the data
6      **for** $k = 1$ to $10$ **do**
7        $testV \leftarrow$ CROSS-VALIDATION(5-fold)
8        $acc(k) \leftarrow$ K-NN$(testV, k)$
9      **end**
10      $[acc\ ind] \leftarrow max(acc)$
11      $k_{best} \leftarrow ind$
12      $acc_{test}(w) \leftarrow$ K-NN$(test, k_{best})$
13    **end**
14    $acc_{avg} \leftarrow mean(acc_{test}(w))$
15    $\sigma \leftarrow std(acc_{test}(w))$
16 **end**
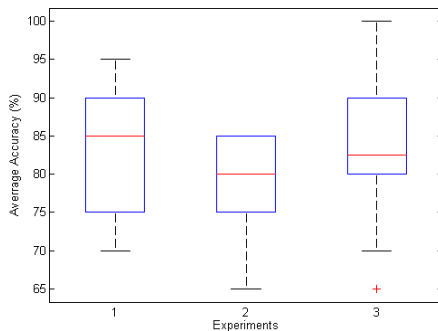17 **retorn** $acc_{avg}, \sigma$

# Experiments and Results

**EX1**: Raw data.
**EX2**: Z-Score Normalization.
**EX3**: Centroid normalization.

## Conclusions

- The BSL recognition is a challenge problem.

- The lack of a database well structured with signs in BSL.

- This paper did an exploratory study of the peculiarities involved in non-manual expressions.

- The methodology adopted had a considerable performance, achieving a maximum average accuracy of 84%.

## Acknowledgment

We would like to thanks PPGEE - UFMG for the incentive and guidance.



The present work was completed with the financial support of CAPES - Brazil.

**Thank you!**