Independent coursework 2

Fingers detection

Course of studies:
Angewandte Informatik

Submitted by: Ilyes Soussi
Matr. -Nr.: s0548859

Supervised by: Prof. Dr. Thomas Jung

University of Applied Sciences Berlin

Contents :

Table of Figures:

Introduction

This project is introduced in the context of the recognition of forms and more specifically in the recognition and interpretation of hand gestures.

In the first part of the report, we will describe the technical and scientific objectives of the project as well as the current used methods in a small bibliographic search.

Then, in a second part, we propose a structural method for the recognition of the gestures of the hand which leads to an interpretation of the method and we will cite the different steps followed.

The last part of the project will be an illustration of the results and the different modules implemented.

Our work focuses on the tracking and calculation of the number of fingers of the hand with a single fixed camera. this aim to provide a low-cost man-machine interface to detect and calculate the number of fingers present via a simple Webcam for training and research purposes within the Faculty of Applied Science Berlin.

I. Scientific description:

1. Technical and scientific objectives:

The recognition of forms and more precisely the gestural study of the hand is one of the approaches that is becoming increasingly important for the naturalization of human-machine interaction.

Many areas rely on pattern recognition. Whether in Biometrics, Virtual Reality, Artificial Intelligence, Augmented Reality, Statistics ... all methods aim at designing a communication interface based on the sensory channel.

The objective of this TP is to develop a method for recognizing and interpreting the gestures of the hand. This TP was based on hand gestures corresponding to figures from one to five.

2. Bibliographic search: Common methods

Different technologies and methods have been developed with the aim of recognizing hand gestures. Below are some of them:

- Data gloves: electrical gloves equipped with sensors providing the position of the hand and the angles of the joints of the fingers
- Gestural interfaces: based on vision through the use of a headset incorporating a camera and a visualization device

The most widely used approaches for interpreting hand gestures can be classified into two types:

- The first approach consists in interpreting the gesture directly from the appearance of the hand in the images, comparing it with a set of model gestures. Appearance models can be based on images, geometric moments, fingertip positions, contour, eigenvectors of images, or a deformable 2d model representing the mean shape of the gesture.
- The second approach is to use a model, which may be voluminal or skeletal. The interest of these models is to offer a very elaborate modelling of the hand, being adjusted to the morphology of the hand, and taking into account the constraints between the joints and the fingers.

II. Proposed method: Skin colour detection

Skin colour analysis is widely used for face and hand detection. The colour of the skin has a characteristic distribution in certain colour spaces, and this property can be used to segment skin colour regions.

There are two important aspects in this type of segmentation: the choice of a colour space and the choice of the method for classifying skin pixels.

1. Hand detection:

The colour space YCbCr, which consists of a luminance component "Y" and two chrominance "Cb" and "Cr", has been used in this project.

It is enough to define explicitly the thresholds of a region of a color space, corresponding to the color of the skin, in our case, as follows:

$$\forall(x,y), \quad S(x,y) = 1 \text{ si } \begin{cases} 77 \leq C_b \leq 127 \\ 133 \leq C_r \leq 173 \end{cases}$$

For this purpose, an image base with hand gestures was used, indicating figures from ONE to FIVE. Below are the results for a gesture with the number "Three":

Figure 1: gesture with the number Three

2. Pre-treatment:

Obviously, the binary image obtained presents false detections and a noise generated by them, which leads us to perform pre-processing on the image in order to detect only the hand in the scene. The operations carried out for this purpose are listed below:

a. Median Filter:

A convolution with a median filter is applied to the image in order to smooth the image and remove the noise around the object.

Median filtering is a nonlinear operation often used in image processing to reduce "salt and pepper" noise. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges.

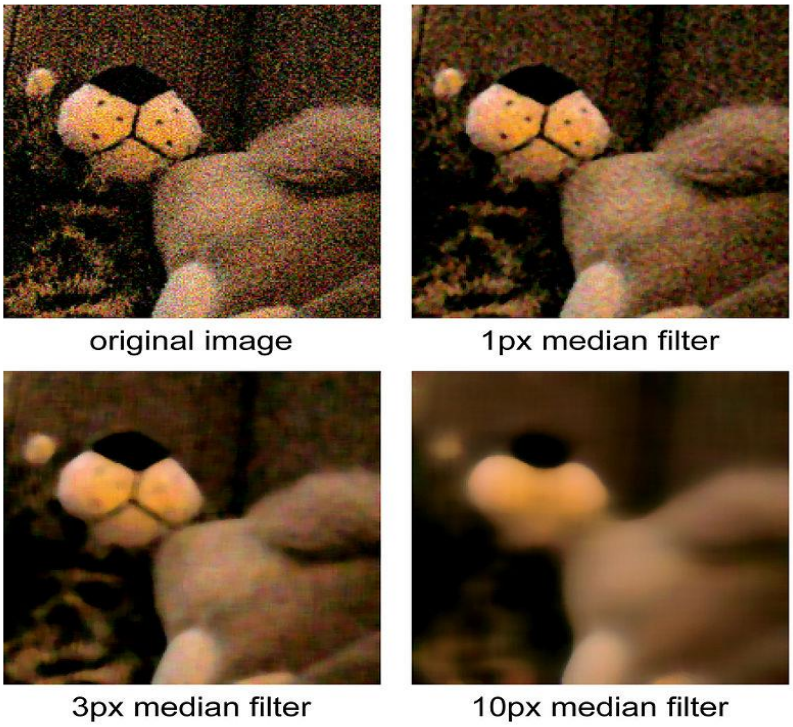

Figure 2: Example of 3 median filters of varying radiuses applied to the same image

b. Closing operation:

Closing is an important operator from the field of mathematical morphology. Like its dual operator opening, it can be derived from the fundamental operations of erosion and dilation. Like those operators, it is normally applied to binary images, although there are gray level versions. Closing is similar in some ways to dilation in that it tends to enlarge the boundaries of foreground (bright) regions in an image (and shrink background colour holes in such regions), but it is less destructive of the original boundary shape. As with other morphological operators, the exact operation is determined by a structuring element. The effect of the operator is to preserve *background* regions that have a similar shape to this structuring element, or that can completely contain the structuring element, while eliminating all other regions of background pixels.

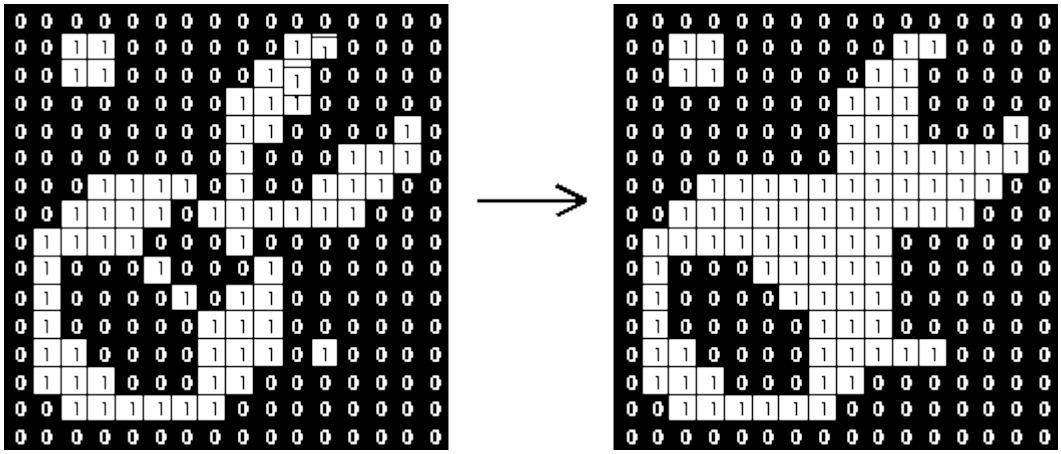

Figure 3: Effect of closing using a 3×3 square structuring element

c. Opening Operation

the opening serves in computer vision and image processing as a basic workhorse of morphological noise removal. Opening removes small objects from the foreground of an image, placing them in the background.

In mathematical morphology, opening is the dilation of the erosion of a set A by a structuring element B:

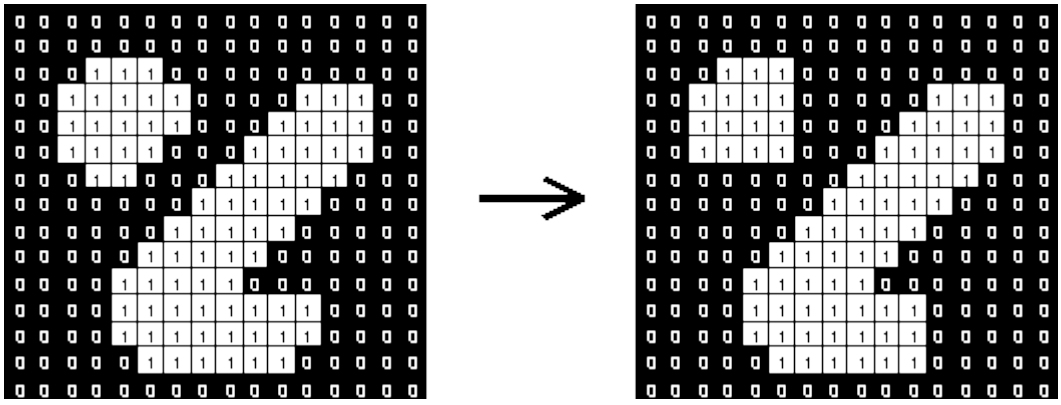

Figure 4: Effect of opening using a 3×3 square structuring element

d. Skeletisation

In order to interpret the gestures of the hand, a Skeletisation of the hand followed by a debarking, which makes it possible to obtain the branches associated with the fingers.

e. Detection of extreme point

The extreme points relative to the ends of the fingers were subsequently detected and counted as indicated by the corresponding gesture.

3. Illustrations

The steps taken during this project allowed us to achieve recognition and interpretation of the hand gestures corresponding to the figures ranging from one to FIVE.

Below are some screenshots of the results:

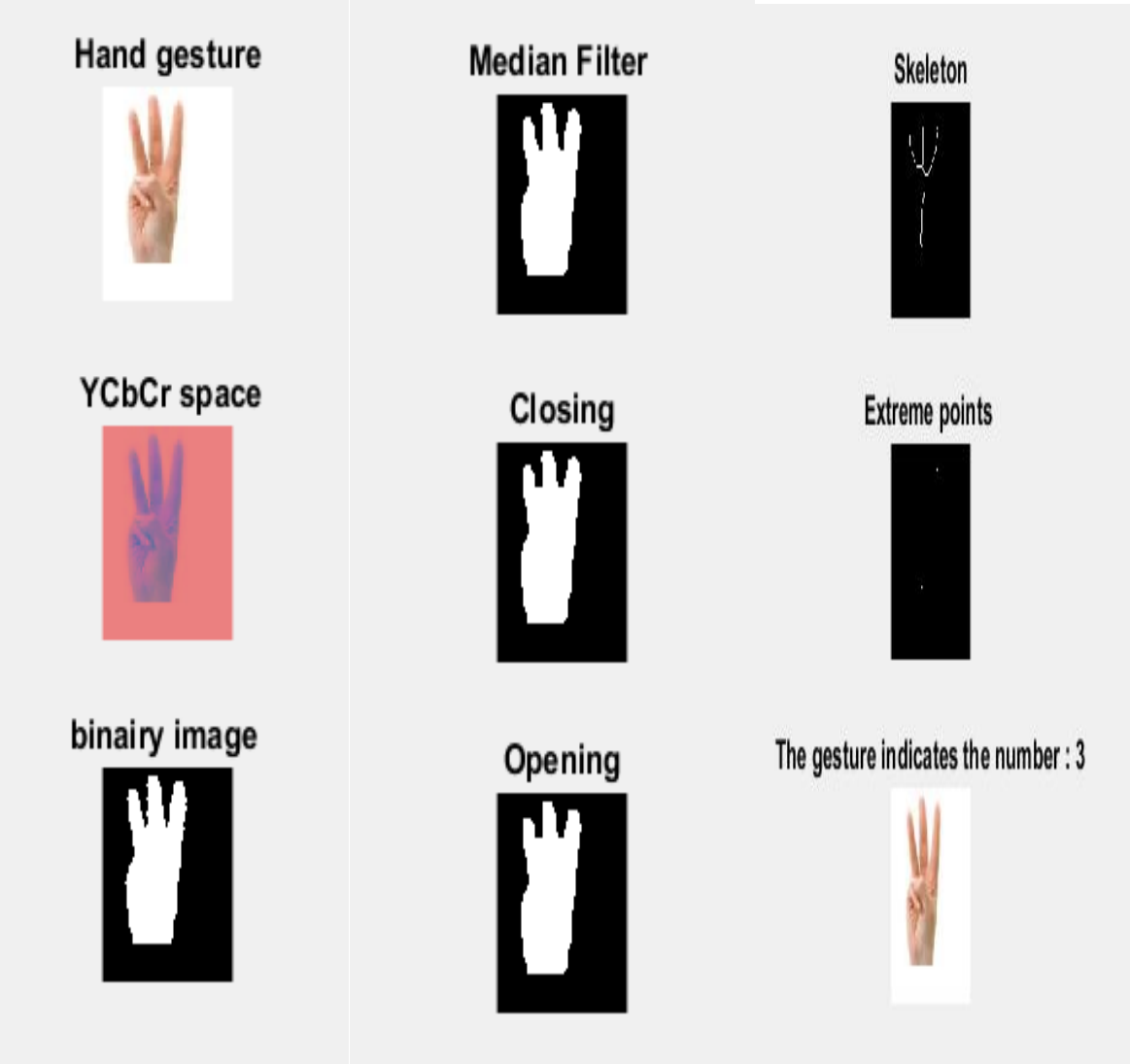

Figure 5: Result of gesture three

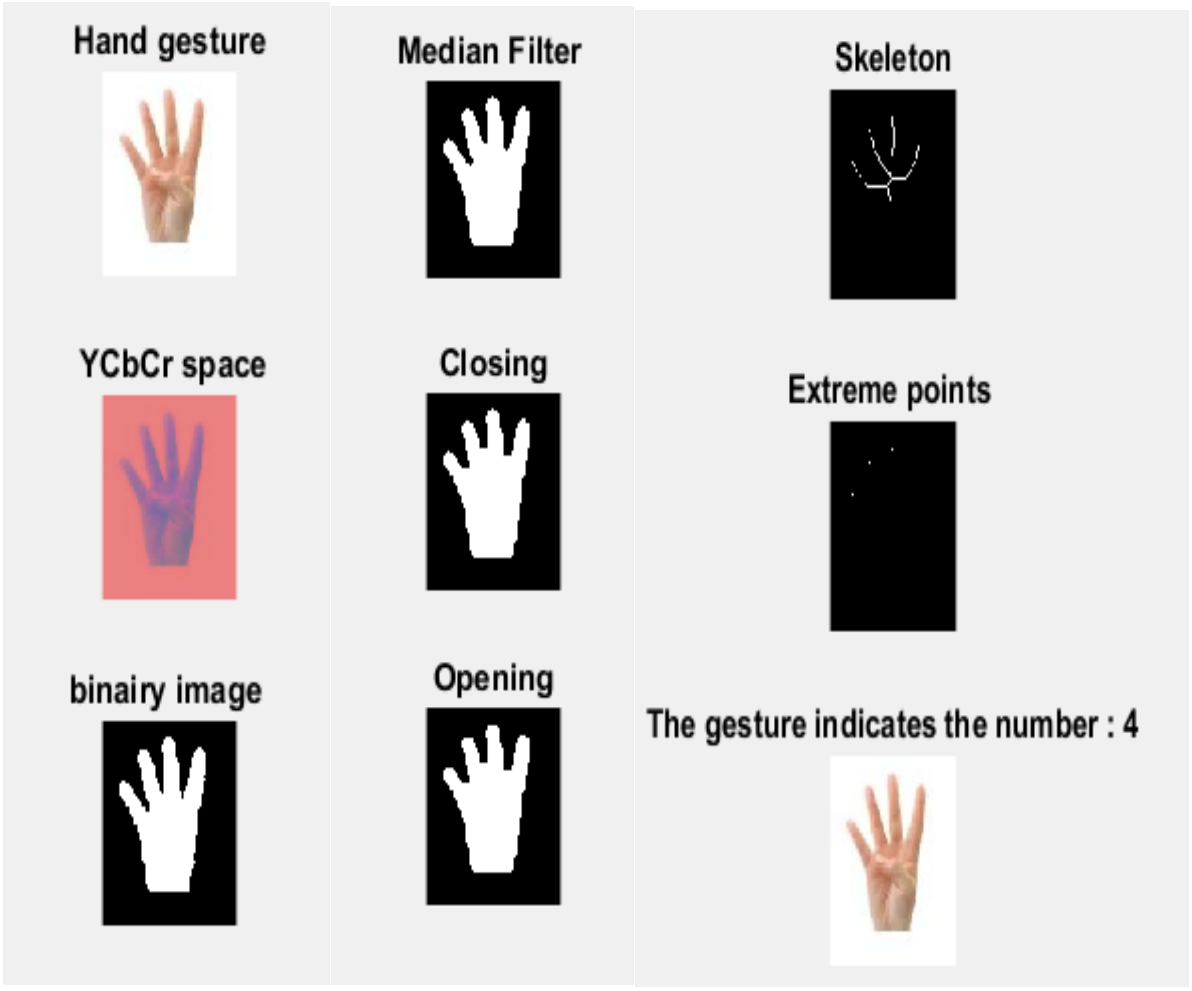

Figure 6: Result of gesture Four

Conclusion

This report presents a structural method leading to the recognition of hand gestures among various other methods.

Thresholding was used to detect the color of the skin followed by pre-treatments to eliminate the artefacts before moving on to a skeletisation that led us back to the interpretation of the gestures.