UNIVERSITÉ CATHOLIQUE DE LOUVAIN

ECOLE POLYTECHNIQUE DE LOUVAIN

Comparison of methods for the number recognition of sport players

Report by:
Gerard Àlvarez Criach

Supervisor:
Christophe De Vleeschouwer

8th of September, 2010

Acknowledgements

It would not have been possible to finish this project without the help and support of some people that I would sincerely want to thank here.

First of all I would like to thank my supervisor Christophe De Vleeschouwer for giving me the opportunity to work with him and especially for all the given help and support.

Also I would like to thank with special affection to François-Olivier Devaux for all the hours he had expended so I could advance my work. And chiefly I wish him good luck in the new stage in his life.

My gratitude also goes to Nicolas Danhier for the help given to me when at first I was completely lost.

I would especially like to thank all my university mates, for all the hours we spent together. It was unforgettable. And I do not want to forget about all the new friends I had met in Louvain-la-Neuve with whom I passed one of the best times of my life.

Finally I dedicate this project and the ending of my university studies to my mother, for her unconditional support and for everything she has done for me. Without her none of this would have been possible.

Gràcies.

Abstract

In a system of detection and recognition of players on a sport-fiel, the process of identifying numbers from a binary image is the last step. The identification is possible using the numbers of players jerseys that are obtained from images captured by a distributed set of cameras. The main problem to solve is that these pictures have different sizes and suffer rotations and non-rigid deformations.

Prior to this project were independently studied three methods to recognize numbers from the players jerseys. The good results obtained encouraged to use any of these methods in the system of recognition of athletes. The provided methods to obtain the image features are based on matching pursuit, local binary pattern and optical character recognition algorithms. The main objective of this project is justify which is the most appropriate method.

But the fact that previous studies were independents involve that can not be made a direct comparison of their results. Therefore is necessary a fair comparison with common criteria. Get this fair comparison is result of working in an homogenous environment. This implies that all the methods have been tested with a set of common images and a common classifier, based on support vector machine. It has been trained with a set of pictures of a basketball match.

The criteria taken into account have been the accuracy in the identification, obtained with the 10-fold cross-validacion, and the time needed by each method to identify a number.

Once done the study, it has been concluded that the method that works with the optical character recognition is the most accurate and fastest. Although this method is the most susceptible to the rotation of images, so its need a normalization, clearly is the most efficient as for the identification of digits as to detect a false identification.

The fact that the differences between the three methods results are so wide, leads to be sure of the choice made. So clearly has been chosen the most efficient method to be used in the project, still in development, that detects, recognizes and tracks players on a sport-field.

Contents

List of Figures

Nomenclature

LBP	Local Binary Pattern
MP	Matching Pursuit
OCR	Optical Chacacter Recognition
SVM	Support Vector Machine

Chapter 1

Introduction

This introductory chapter includes an overall overview of the proposed problem to be solved, the tools used and a brief explanation of the procedures to determine the solution. It is also included an explanation of the framework where this work will be enclosed and finaly it is outlined the rest of the paper.

1.1 Overview

To achieve an objective there are many possible ways, each one with its advantages and disadvantages. The main aim of this project is to discover the most efficient path to realize this goal, which is recognize a number from a binary image. The different ways are the identification methods listed later.

But, which can be the use for the recognition of numbers? One use is the identification of athletes using the images provided by a network of cameras placed in the sport-field. This application was developed by Nicolas Danhier[4], in collaboration with the UCL and the European FP7 Apidis project[1], and it achieves the detection and recognition of basketball players using the jersey number.

The Danhier's system, outlined in Figure 1.1, detects the situation of each player. Detection relies on the definition of an occupacion probability map of the ground. In these positions, cuts out a bounding box of 2 meters high by 80 centimeters wide. Within the process of detection the positions, the team each player belongs is identified[6]. The bounding box image is segmented, using a mean shift algorithm, and the digit region is selected. The selection is based on contextual attributes as that the valid digit regions never touch the border of the bounding box or that are surrounded by a single homogeneously colored region. Once detected the digit in the bounding box picture, a binary image is generated with the digit shape and characterized by a vector of features, based

[1] Autonomous production of images based on distributed and intelligent sensing (www.apidis.org)

on the optical character recognition, for a later identification using a support vector machine classifier.

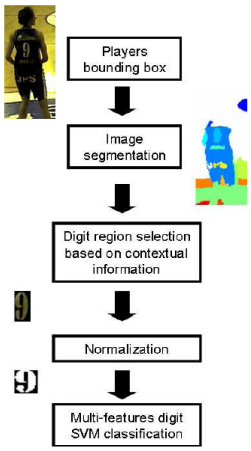

Figure 1.1: Recognition of digits printed on players shirts through segmentation, selection, and classification of regions that are likely to represent digits.

Parallel to this project, UCL staff has tested independently different manners to achieve the recognition of numbers from binary images. So the three provided methods to compair their efficiency are based on:

- ▷ Optical character recognition [5]
- ▷ Local binary pattern [1]
- ▷ Matching pursuit [2]

Any of these methods can work in the last stage of the players recognition, but the fact that these studies have been independent, with different criteria, does not allow a direct comparison of the previous results. Therefore it is necessary an environment with the same conditions for a proper comparison. This is about the project has focused, make a fair comparison between them.

To perform this study it has been used a set of images of a basketball match, provided by Apidis[2]. This dataset contains color pictures where are located clearly the players numbers, that were obtained by the mean shift segmentation of the bounding box pictures. Each one has followed a process to isolate the digit in a binary image sized at 32 by 32 pixels, based on Danhier's work. These images will be characterized by each of the methods.

[2]The dataset is publicly available at http://www.apidis.org/Dataset/

The final stage in the recognition compares the obtained vector of features with a set of training vectors to allow a subsequent classification. This process is performed using a support vector machine classifier, choosen instead other classifiers based on minimum mean distance or k-nearest neighbours.

Achieved an homogeneous environment, using a common set of images for all the study and the same classifier for each identification procedure, it only remains establish the quality of each method. The criteria to find the most efficient system have been:

- ▷ The balanced classification rate (%)
- ▷ The delay needed to the identification (seconds)

The percentage of correct recognitions is found using the 10-fold cross-validation. To calculate the recognition time required, it is done a mean of the time needed to identify each sample.

The result of this work will determine whether the currently method used by the Apidis system is the most efficient or if any of the other proposed procedures would be more appropriate.

1.2 State of the art of detecting and recognizing players on a sport-field

The Apidis project is working on a framework to automate the collection and distribution of digital contents, specially to automate the production of video contents for specific scenarios, as sports events or surveillance. In the long term, the system will cover local sporting events without needing technical employers. By analyzing and interpreting images, the device will be able to automatically generate summaries of the matches, provide player statistics...

The semantic understanding of sports video can be considered as understanding the process and strategy of the game. The position and trajectory of players in the game can provide useful information to facilitate such semantic analysis. The detection and tracking reveals the movement of the players on the field, which can be used to detect certain events or understand the overall trend and strategy of the game.

But Apidis is not the only entity that works on the recognition and tracking of athletes. Iwase and Saito [11] propose a method to track soccer players using multiple views, that performs inner-camera operations to track the players and, in any case that players can not be tracked, uses inter-camera operation. But players are tracked with initial information, the index numbers, in the initial frame of each camera.

The Harbin Institute of Technology of China[9] proposes an alternative method for athlete identification. Firstly, image segmentation, based on generalized learning vector quantization algorithm, is employed to separate the jersey number regions of its background. And size/pipe-like attributes of digital characters are used to filter out candidates. Then, a K-nearest neighbor classifier

is employed to classify a candidate. In the recognition procedure, they use the Zernike moment features, which are invariant to rotation and scale for digital shape recognition. Synthetic training samples are used to represent the pattern of digital characters with non-rigid deformation.

Finally, another system developed by Harbin Intitute of Technology [7] too, allows the detection and tracking approach based on support vector machine and particle filter to track players in broadcast sports video.

1.3 Outline

The rest of this paper is organized as follows. The next chapter will introduce the three methods used to characterize images. Chapter 3 focuses on explaining the tools and algorithms used for the comparison. Experimental results are presented in chapter 4 and finally the results interpretation and the conclusions are shown in the last chapter.

Chapter 2

The studied methods

This chapter will present the three studied methods to characterize images. A crude representation of information often leads to large vectors. The aim is to compress information in a minimum discriminatory features to reduce the complexity of the calculations of the algorithms. Each of these methods provides a vector of features that will be used to classify and identify the input image. The way how are obtained is what this section focuses.

2.1 Introduction

All character identification systems are based on shape recognition from binary images. The text detection or the identification of vehicles by the number plate are some uses. But all applications get the pictures from non-binary images where normally are not isolated the shapes to reconize. The captured images undergo a segmentation process. This refers to the process of partitioning a digital image into multiple segments. More precisely, it assigns a label to every pixel in an image such that pixels with the same label share certain visual characteristics. So is able to locate objects and boundaries in pictures. But this implies that the shape of the object to be isolated from the original image is kept.

In applications already in use, the images are taken from fixed objects or with known orientation. This means that the inclination of the characters gonna be vertical, with no distortion and no hidden. But the use of the proposed identification methods will be the recognition of sport athletes by the jersey number. The digits are located on the players shirts of players who make unpredictable movements. Also the cloth movements deform the figures shapes. That means that the main problems come up from the orientation, size and non-rigid deformations of the digital characters and the possibles obstructions.

The system designed by Danhier [5] solves many of these problems. Working with a multi-view player detection solves the problem that numbers can be partially or completely covered. If the number is not visible for a camera can

be taken by any of the others. Moreover in the step to obtain the binary image from the number, the deformation effects are minimized and the system allows a rotation normalization of the digits. That implies horizontal alignment of the major principal axis. This is the context where will be used the three methods introduced below.

Each method presentation will have a short introduction and a detailed explanation about how are characterized the images.

2.2 Matching pursuit

The algorithm called MP decomposes any signal into a linear expansion of waveforms that belong to a redundant dictionary of functions. Such functions are called time-frequency atoms.

Such decompositions are similar to a text written with a small vocabulary. Although this vocabulary might be sufficient to express all ideas, it requires tu use alternatives that replace unavailable words by full sentences.

The work of [10] proposes to use a MP algorithm to generate image representations over a redundant dictionary, which is specifically designed to capture 2-D features of natural images. The dictionary is built by anisotropic refinement and orientation of contour-like functions to capture edges.

Given a d-dimensional signal f in a real vector space, the central problem is to compute a good approximation $\tilde{f_N}$ as a linear superposition of N basic elements, which are often called atoms, picked up in a huge collection of signals D, usually referred to as a dictionary. The dictionary is said to be redundant when $D \gg d$. The approximant $\tilde{f_N}$ is sparse when $N \ll d$, where

$$\tilde{f_N} = \sum_{k=0}^{N-1} c_k g_k \ , \ g_k \epsilon D.$$

There are no particular requirements concerning the dictionary, except that it should span the signal space H, and there is no prescription on how to compute the c_ks.

The dictionary is composed of atoms that are built on Gaussian functions along one direction, and on second derivative of Gaussian functions in the orthogonal direction.

$$g\left(\overrightarrow{p}\right) = \frac{2}{\sqrt{3\pi}}\left(4x^2 - 2\right)\exp\left(-\left(x^2 + y^2\right)\right)$$

Where $\overrightarrow{p} = [x, y]$ is the vector of discrete image coordinates and $\|g\left(\overrightarrow{p}\right)\| = 1$.

The generating functions described above are, however, not able to efficiently represent the low frequency characteristics of images at low rate. To capture these features it is used an additional dictionary capable of representing the low frequency components. The second generating function of our dictionary can be written as:

$$g(\overrightarrow{p}) = \frac{1}{\sqrt{\pi}} \exp - \left(x^2 + y^2\right)$$

Where the Gaussian are mulyiplied by a constant in order to have $\|g(\overrightarrow{p})\| = 1$.

To put it more simply, the decomposition of images to classify is based on: instead of describing the images in the pixel domain, has been tested to describe them with elongated ellipses in order to recover the strokes that make up the number. This decomposition is possible by the MP code [1]. An example of reconstruction of grayscale images is shown in Figure 2.1.

Some examples of decomposition used in this project are shown in Figure 2.2. Can be seen that using the first six ellipses is able to recognize any of the numbers.

Figure 2.1: Example of image reconstruction using the MP algorithm with few atoms.

Figure 2.2: Original image and reconstruction with six atoms based on the MP algorithm.

[1] The algorithm is publicly available at http://lts2www.ep.ch/jost/-downloads/TBP2D1.11/index.html

2.2.1 Features used

Students that previously worked with this method[2] used the six first ellipses of the MP. Once the image is represented with ellipses, they calculated three features per atom:

- ▷ The distances between the global center of gravity of the whole figure (g) and the centers of the ellipses.
- ▷ The angle α_1 between the straight lines connecting the centers of the ellipses to g.
- ▷ The angle α_2 between the major axis of the ellipses with the line that connect their centers with g.

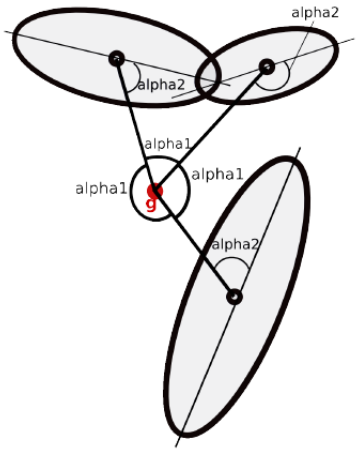

Figure 2.3: Representation of MP features with 3 ellipses.

The choice of these features was motivated by the desire to take robust features to rotation. The figure 2.3 shows an example of the listed features using three ellipses.

The calculation of the variables was obtained from the positions of the centers of the ellipses and their orientation:

- ▷ The distances to g: $\sqrt{\left(p_x\left(i\right)-p_x\left(g\right)\right)^2+\left(p_y\left(i\right)-p_y\left(g\right)\right)^2}$ where p_{x} and p_{y} are respectively the x and y positions of the i-th ellipse and the center.
- ▷ $\alpha_2 = \arccos\left(\left(\left[\cos\left(angl\left(i\right)\right),\sin\left(angl\left(i\right)\right)\right]*\left[dx\left(i\right),dy\left(i\right)\right]'\right)/d\left(i\right)\right)$ where angl(i) is the orientation of the i-th atom in relation to the vertical, dx and dy are the differences between g and i, d(i) is the distance between i and g and $'$ represents the transpose of the vector.

▷ $\alpha_1 = \arccos\left(\left(d(i)^2 + d(j)^2 - d(i,j)^2\right) / (2 * d(i) * d(j))\right)$ where d(i,j) is the distance between i and j, the rest of notation is the same as for α_2. This formula is an application of Theorem d 'Al Khashi (Pythagoras generalized).

2.3 Local binary pattern

This section introduces the definition of a descriptor based on LBP features distribution. To develop this descriptor the previous study was based on the work of [12].

The LBP operator was originally designed for texture description. The operator assigns a label to every pixel of an image by thresholding the 3x3-neighborhood of each pixel with the center pixel value and considering the result as a binary number. Then the histogram of the labels can be used as a texture descriptor. See Figure 2.4 for an illustration of the basic LBP operator.

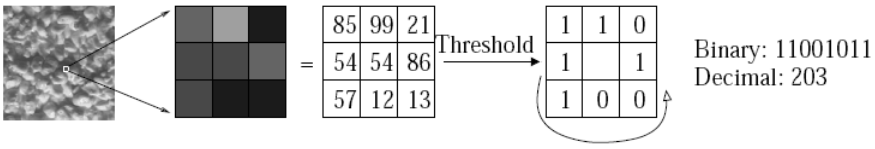

Figure 2.4: The basic LBP operator.

To be able to deal with textures at different scales, the LBP operator was later extended to use neighborhoods of different sizes. Defining the local neighborhood as a set of sampling points evenly spaced on a circle centered at the pixel to be labeled allows any radius and number of sampling points. Bilinear interpolation is used when a sampling point does not fall in the center of a pixel. In the following, the notation R and n will be used for pixel neighborhoods which means n sampling points on a circle of radius of radios R. See Figure 2.5 for an example of circular neighborhoods.

2.3.1 Features used

LBP operator is applied to a pixel and takes as arguments: a radius R, a number of points n and a label is assigned to this pixel. The label is the number of 1 - 0 or 0 - 1 transitions found when are covered "n" points equally spaced around the circle of radius R centered on the pixel. The value of a point is the value of the pixel where is located. If the point is located outside the image, the assigned value is the color of the background. The number of possible labels is $\left[\frac{n}{2}\right] + 1$.

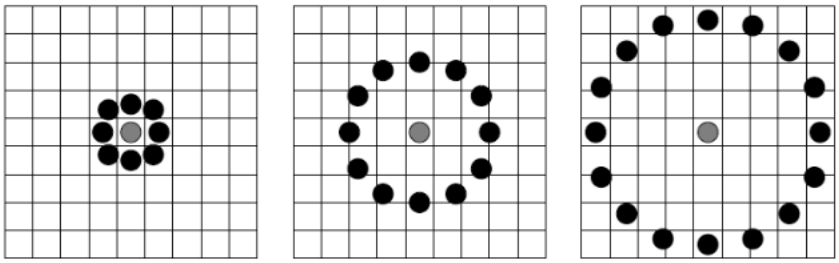

Figure 2.5: Circularly symmetric neighbor sets. The pixel values are bilinearly interpolated whenever the sampling point is not in the center of a pixel.

To the descriptor calculation, the label of each pixel is established. There are two types of labels: those assigned to a white pixel (value 1) and assigned to a black pixel (value 0). An histogram is created to count the number of cases of each label type for each label separately. The descriptor is a vector showing this histogram.

In many applications, specifically facial recognition, the image is divided into several regions and an histogram is calculated for all the different regions. In this case has been chosen to use a global histogram because otherwise, the descriptor is sensitive to the location and orientation of the number.

2.4 Optical character recognition

This is the method to characterize images used by the Danhier's system[5]. It was based on the many suited features used to character recognition, for example described in [8]. In practice, the OCR classifier provides vectors with the following features.

2.4.1 Features used

This method provides a vector with 38 parameters that are allocated as follows:

- ▷ 2 parameters: The x-y vector that specifies the center of mass of the region.
- ▷ 1 parameter: The "EulerNumber", a scalar that specifies the number of objects in the region minus the number of holes in those objects.
- ▷ 3 parameters: The image moments m_{02}, m_{20} and m_{22} that are calculated with the formula:

$$m_{pq} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - x_G)^p (y - y_G)^q I(x,y)}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y)}$$

M=N=32, I represents the binary image and (x_g,y_g) is the gravity center position.

▷ 2 * 16 parameters: The projections in the horizontal and vertical directions. In the Figure 2.6 are graphical representations of these projections.

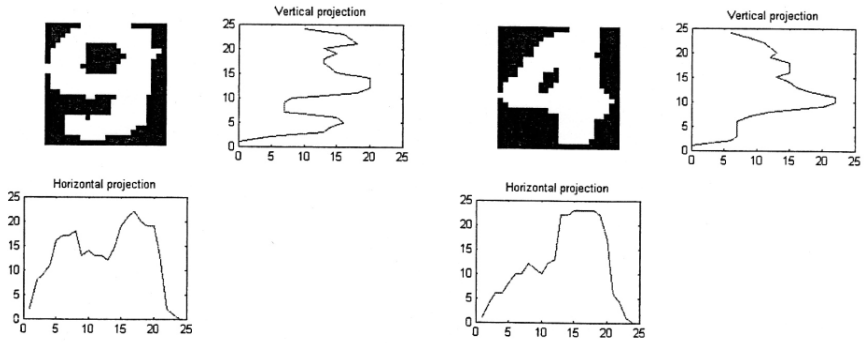

Figure 2.6: Example of horizontal and vertical projection.

2.5 Conclusions

Apparently all these procedures to characterize images are valid enough. They were chosen because previous studies concluded that their use results in the identification of images were correct. But it makes no sense to draw a conclusion from these results because all these studies were carried out according to different criteria. So it makes sense to carry out a proper and fair comparison.

Importantly, the features of the OCR method are susceptible to rotation, so the input images require a rotation normalization. However the other methods do not, because they have concluded that are robust to rotation. This implies an algorithm efficiency, also assessed in the final result of the comparison.

Chapter 3

Test protocol

This chapter will show the methodology used for the comparison of the different methods, the utilized tools and the developed algorithms.

3.1 Introduction

A pattern recognition system aims to predict which class belongs an unknown shape. In this usage, the shapes to recognize are the obtained images from the segmentation of the basketball player pictures, and the labels which must be classified are all the players numbers of each team.

To this end some steps are required regardless of method used, outlined in Figure 3.1. Firstly, the identification system needs an input signal. This is the binary image with the shape to identify. Then the image is characterized by any of the provided methods, obtaining a vector of features. This vector is entered to a classifier, previously trained, that will provide the predicted label.

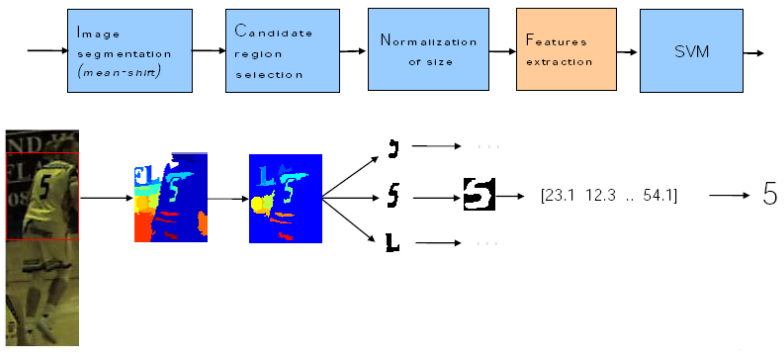

Figure 3.1: Structure of the number recognition system. The underlined step is where can work any of the three studied methods.

To perform this comparison is used a set of images containing the numbers of basketball players. Through these images, the chosen classifier is trained. Once the classifier is ready, it is possible to find the parameters that will show the quality of each tested method. Show these tools is what focuses this chapter: the used dataset, the classification algorithm and the criteria to find each method efficiency.

3.2 Set of images

The identification system of basketball players, developed by Danhier[4], locate the position of the players through a network of video cameras. This position is used as an anchor for the player bounding box projected in each one of the views, as shows Figure 3.2. Within this bounding box, the regions that are like a digit are selected, provided by mean-shift segmentation.

Figure 3.2: Examples of some bounding box pictures.

The set of images used to test the system comes from the images obtained by the bounding box segmentation. Some of the pictures used are shown in Figure 3.3. Importantly, the images have different sizes because they depend on the distance between the basketball players and cameras.

Figure 3.3: Examples of the training base.

The dataset contains 2939 samples distributed as follows:

- ▷ Team A (yellow):
  - 421 numbers "4"
  - 130 numbers "5"
  - 335 numbers "8"
  - 350 numbers "9"
  - 150 numbers "10"
  - 88 numbers "11"
  - 175 numbers "14"
- ▷ Team B (black)
  - 234 numbers "4"
  - 37 numbers "6"
  - 275 numbers "7"
  - 23 numbers "8"
  - 101 numbers "9"
  - 149 numbers "11"
  - 287 numbers "13"
  - 184 numbers "15"

But it is not advisable use directly the regions resulting from segmentation because are rough. The main of this stage is isolate the shape to recognize and prepare it to the statistical processing.

Firstly is necessary convert the input image, from grayscale to binary, using a global threshold found with Otsu's method (available in Matlab). Then is detected the digit region and cut that position of the input RGB picture. At this point, images follow a different process depending on whether they need to normalize their rotation. Normalization implies horizontal alignment of the major principal axis, as derived through computation of moments of inertia. After correcting the rotation, if it is required, the images follow the same process as above: binarization, re-locate the digit area and cut that region of the input RGB image. Importantly, if the rotation has been normalized, the last crop is made at the same angle. Finally is obtained an image with the digit isolated. The last step is to resize to 32 by 32 pixels and binarized image. Once the binary mask is done, it is important to ensure that the number is represented by white pixels and black background. This whole process is outlined in Figure 3.4.

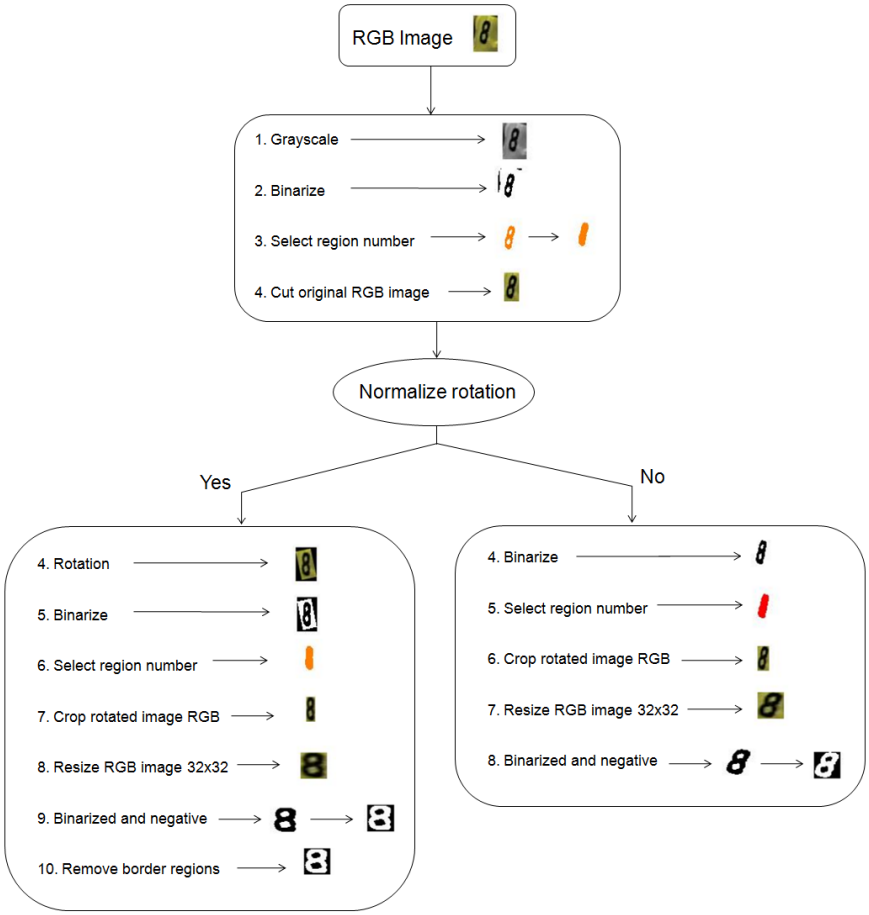

Figure 3.4: Pretreatment process of the RGB images.

Figure 3.5 and 3.6 show an example set of samples, the first one without the normalization of rotation and the other normalized.

But the system must be capable to make out a figure of a false detection. For this reason will be introduced a new label composed by false detections generated with the segmentation of pictures where there are no digits (Figure 3.7). This label contains 1230 samples. Predictably the recognition rate will decrease.

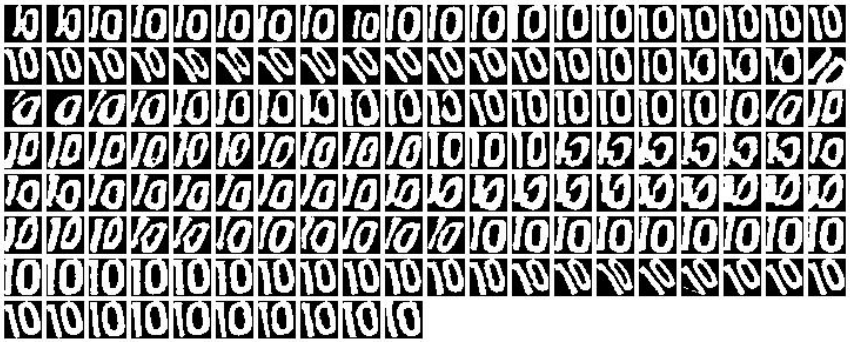

Figure 3.5: Display of the number 10's samples without de rotation normalized.

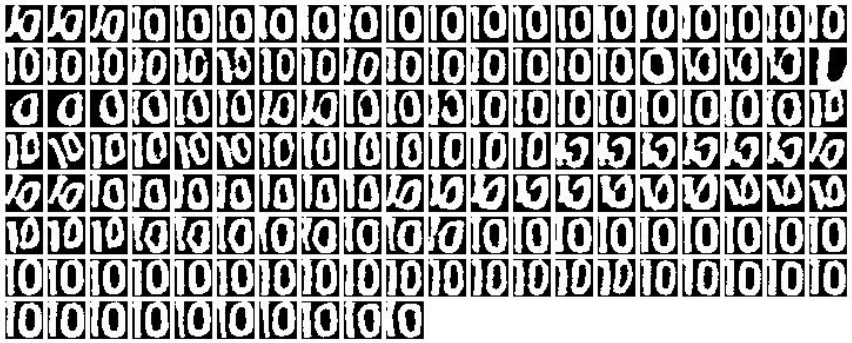

Figure 3.6: Display of the number 10's samples with de rotation normalized.

3.3 The classification algorithm

The final stage of pattern recognition is to compare the vector obtained from the features of the studied shape, with vectors of the training base. In this project it is chosen a SVM classifier, more precisely the LIBSVM library [3].

SVM has been successfully applied to a wide range of pattern recognition and classification problems. The advantages of SVM over other methods consist of:

1. Providing better prediction on unseen test data.
2. Providing a unique optimal solution for training problem
3. Containing fewer parameters compared with other methods.

A classification task usually involves separating data into training and testing sets. Each instance in the training set contains one "target value" (the class labels) and several "attributes" (the features). The goal of SVM is to produce a model (based on the training data) which predicts the target values of the test

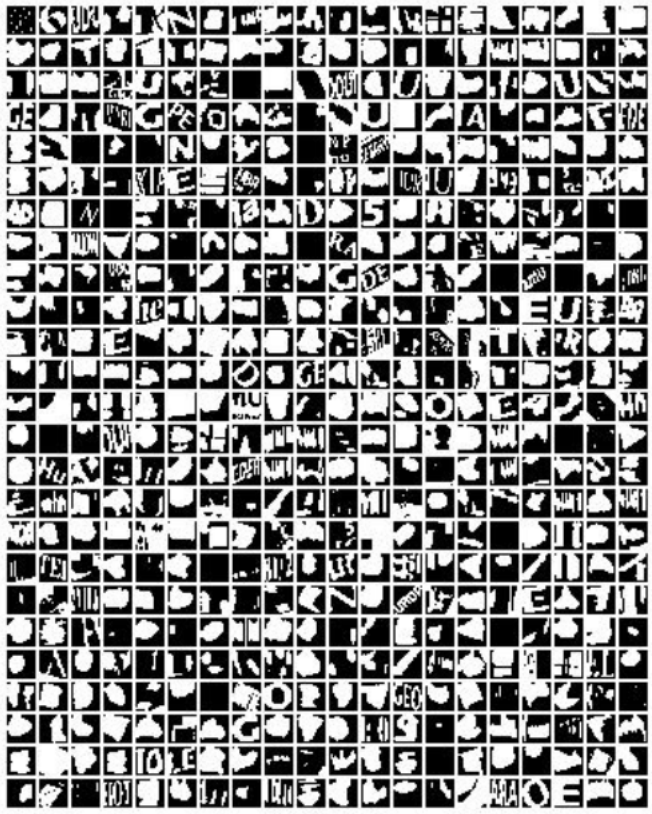

Figure 3.7: Partial display of the noise label.

data given only the test data attributes. SVM requires that each data instance is represented as a vector of real numbers. Training vectors are mapped into a higher dimensional space by the function ϕ. SVM finds a separating hyperplane with the maximal margin in this higher dimensional space. C > 0 is the penalty parameter of the error term. Furthermore, $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ is called the kernel function. Though new kernels are being proposed by researchers as the following four basic kernels:

- ▷ Linear: $K(x_i, x_j) = {x_i}^T x_j$
- ▷ Polynomial: $K(x_i, x_j) = \left[\left(\gamma (x_i)^T\right)(x_j) + r\right]^d, \gamma > 0$
- ▷ Radial basis function (RBF): $K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0$
- ▷ Sigmoid: $K(x_i, x_j) = \tanh\left(\gamma (x_i)^T (x_j) + r\right)$

Here, γ, r, and d are kernel parameters.

The RBF kernel is the chosen. This kernel nonlinearly maps samples into a higher dimensional space so it, unlike the linear kernel, can handle the case when the relation between class labels and attributes is nonlinear.

The output of the recognition function of players is simply the output of the classifier, that provides the predicted label.

3.4 Evaluation criteria

It only remains to decide the parameters for evaluating the tested methods. Clearly accuracy is the most important parameter to be taken into account. But it has also assessed the time needed by each method to identify the input image.

3.4.1 10-fold cross-validation

As discussed above, a common strategy is to separate the dataset into two parts, of which one is considered unknown. The prediction accuracy obtained from the "unknown" set more precisely reflects the performance on classifying an independent dataset. An improved version of this procedure is known as cross-validation.

In v-fold cross-validation, the training set is divided into "v" subsets of equal size. Sequentially one subset is tested using the classfier trained on the remaining "v-1" subsets. A simple scheme is shown in Figure 3.8. Thus, each instance of the whole training set is predicted once so the cross-validation accuracy is the percentage of data which are correctly classified.

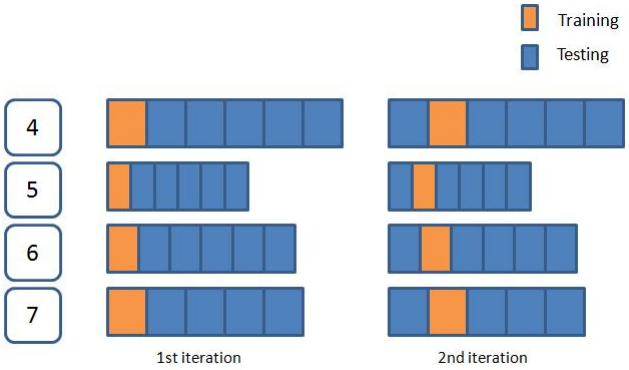

Figure 3.8: v-fold crossvalidacion scheme, showing the division of the dataset in "v" equal parts for each of the labels.

In this case it is used a 10-fold crosvalidacion. Each label is divided into subsets of 10% of total samples and therefore have taken 10 iterations. The accuracy of the tested method is the average of each iteration accuracy.

3.4.2 Identification delay

It is also interesting to know the time required by each method to identify a digit. That means the time required to extract the features of the input image and compare them with the model generated by the SVM.

It has been calculated the average delay needed for each sample identification.

3.5 Conclusions

Finally has been established a homogeneous environment for a fair comparison. All methods are tested with the same set of images, each one works with the same SVM classifier and will be evaluated with the same criteria. Just remains evaluate the results.

Chapter 4

Experimental results

Below are all the test results. Firstly it is interesting to repeat the tests of previous studies, trying to find the same results. To accomplish this, it has been used the set of images of the previous studies but with the actually developed algorithms. This way will check the proper functioning of the developed systems. Each section shows the new test results of each method too. Furthermore all the methods are tested firstly without a false detection label and then using it. Finally the results are compared to conclude the best procedure.

4.1 Matching pursuit

First of all are introduced the tests conducted in the previous study of the MP method [2]. Moreover these tests are repeated using the same set of images that were used previously, but with the current SVM classifier. In this way will be checked whether the developed algorithms are done correctly.

4.1.1 Repetition of the MP's previous results

To evaluate the performance of the MP, first was made a binary classification of numbers as shows the Figure 4.1. All tests were made on cross-validation, obtaining the classification rate. First are shown the results of the previous study and then the repetition of the test to check the correct performance.

Then was tested a multi-class classification, first separating the digits of each team. The classification of all team A numbers achieved an accuracy of 85.65%. For the team B the precision was 90.22%. The Figure 4.2 shows the confusion matrix for each team.

Finaly it was tested the classifier with the whole dataset getting this confusion matrix and obtaining a 81.60% accuracy. The Figure 4.3 shows the confusion matrix with the totality of labels of each team.

Team	Num	Team	Num	Previous accuracy	New accuracy
A	4	B	6	94.73	99.27
A	4	B	7	83.44	98.18
A	4	B	15	91.08	99.61
A	4	A	14	88.04	99.23
A	10	A	14	87.05	98.46
A	5	B	6	91.95	95.88
A	5	B	15	88.29	93.80
A	9	B	4	93.94	97.97
A	9	A	4	93.38	97.53

Figure 4.1: Number identification accuracy using only two labels.

A	4	5	9	10	14
4	18	0	1	1	2
5	1	5	1	2	0
9	0	1	46	0	0
10	0	2	1	9	0
14	1	0	0	0	7

B	4	6	7	9	15
4	14	0	2	0	1
6	0	1	1	1	0
7	0	0	19	0	0
9	0	0	0	9	0
15	0	0	0	0	8

Figure 4.2: Confusion matrix of the reconized samples of Team A and Team B.

The repetitions of the previous tests show some differences in the results, as shows the Figure 4.4. As can be seen in Figure 4.3, the previous study of the MP's method did all tests with a limited number of samples of each number. For example, for the number 4 of the team A, they had been used 22 samples. Given the impossibility of knowing which samples they used, the tests have been repeated with all the samples of their dataset. This may explain the differences in the results.

	4	5	9	10	14	4	6	7	9	15
4	29	1	0	0	1	1	0	0	0	0
5	2	13	1	1	0	0	0	0	1	0
9	0	0	28	2	0	0	0	0	0	0
10	0	0	3	5	0	0	0	0	1	0
14	1	0	0	1	15	0	0	0	0	0
4	0	0	0	0	0	8	0	0	0	0
6	0	0	0	0	0	0	3	0	0	0
7	0	0	0	0	0	0	0	15	0	1
9	0	0	1	0	0	0	0	1	7	1
15	0	0	0	0	0	0	0	1	0	10

Figure 4.3: Confusion matrix with samples of Team A and Team B.

	Previous accuracy	New accuracy
Team A numbers	85.65	67.72
Team B numbers	90.22	59.50
Totality of numbers	81.60	63.50

Figure 4.4: Comparation between the previous results and the repetition of the tests.

4.1.2 The new tests of MP's method

Below are shown the new tests of this method. It has been used the set of images introduced in Section 3.2, without the normalization of the rotation. Are also used all the tools and criteria explained in Chapter 3.

Recognition rate numbers without noise label

Before testing whether the system is robust to detection, it is important to check if the numbers are differenced between them. It is interesting to introduce the confusion matrix of this method in Figure 4.5. As can be seen it tends to accumulate on the diagonal the peak values.

The Figure 4.6 shows the histogram of the numbers 8 and 9, where can be seen that none of the columns clearly discriminates the label. The histogram of the numbers 7 and 14, Figure 4.7, shows various components of the feature vector that discerns the labels, as the 9th column.

Once it is shown the performance of this method, the classification rate obtained, using only the digit labels, is:

55.51 %

1.	4	5	6	7	8	9	10	11	13	14	15
4	583	1	0	3	44	27	0	17	0	1	3
5	12	15	0	34	6	37	1	1	15	7	2
6	6	0	1	1	5	4	5	2	10	0	4
7	16	10	0	199	5	14	2	15	7	0	5
8	56	0	0	5	150	60	7	18	8	41	9
9	51	0	0	8	24	315	10	1	5	35	3
10	20	3	0	9	23	34	13	11	23	0	14
11	8	2	0	5	12	57	3	61	60	26	3
13	3	2	0	2	2	39	1	21	186	7	20
14	10	0	0	0	2	52	0	0	0	109	0
15	7	2	0	13	9	33	14	4	73	5	24

2.	4	5	6	7	8	9	10	11	13	14	15
4	85.86	0.15	-	0.44	6.48	3.98	-	2.50	-	0.15	0.44
5	9.23	11.54	-	26.15	4.62	28.46	0.77	0.77	11.54	5.38	1.54
6	16.22	-	2.70	2.70	13.51	37.84	5.41	40.54	18.92	-	13.51
7	5.86	3.66	-	72.89	1.83	5.13	0.73	5.49	2.56	-	1.83
8	15.82	-	-	1.41	42.37	16.95	1.98	5.08	2.26	11.58	2.54
9	11.28	-	-	1.77	5.31	69.69	2.21	0.22	1.11	7.74	0.66
10	13.33	2.00	-	6.00	15.33	22.67	8.67	7.33	15.33	-	9.33
11	3.38	0.84	-	2.11	5.06	24.05	1.27	25.74	25.32	2.95	1.27
13	1.06	0.71	-	0.71	0.71	13.78	0.35	7.42	65.72	2.47	7.07
14	5.78	-	-	-	1.16	30.06	-	-	-	63.01	-
15	3.80	1.09	-	7.07	4.89	17.93	7.61	2.17	39.67	2.73	13.04

Figure 4.5: MP's confusion matrix. The first table cointains the matrix using the number of samples. The second table contains the percentages (%).

Recognition rate numbers with noise label

The accuracy obtained is quite low, so it might be concluded that this method is not efficient in the recognition of numbers. But the testing protocol will be followed. If the false detection label is used, the classification rate is:

30.01 %

As was expected the recognition rate decreases, obtaining a value too low to be used in a players recognition system.

Identification delay

The mean made from the times needed to identify each of the samples gives the following average value of the delay:

3.6398 seconds

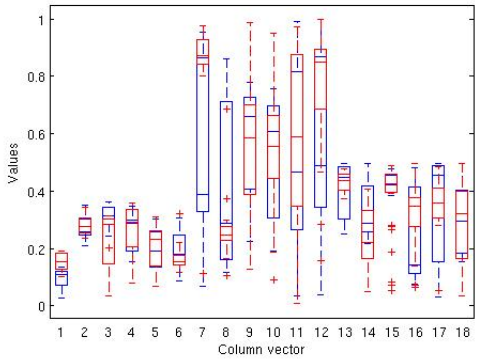

Figure 4.6: Histogram of the MP's samples of the numbers 8 (blue) and 9 (red) with R=8 and n=8.

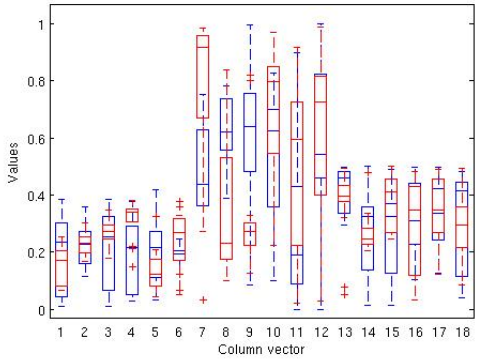

Figure 4.7: Histogram of the MP's samples of the numbers 7 (blue) and 14 (red) with R=8 and n=8.

4.1.3 MP's conclusions

It has been shown that the use of MP-based features is not suitable for a number identification system. Using a multiple classifier, the identification accuracy decreases as the number of labels is increased. The classification rate is too low and the recognition delay is too wide.

4.2 Local binary pattern

Firstly this section will show the tests carried out in the previous study of the LBP method [1]. Furthermore these tests are repeated using the same set of images of the previous study, but with the current SVM classifier. This way will check whether the developed algorithms are done correctly.

4.2.1 Repetition of the LBP's previous results

To test the discriminative ability of the descriptor, were used 200 samples of the numbers 4, 7, 9 and all available samples of the numbers 5, 10, 14, 15 in order to obtain a balanced classification rate of all tested labels.

The Figure 4.8.1 shows the classification rate obtained with all the introduced labels (4, 5, 7, 9, 10, 14, 15). The classification rates were better the more points n were used. As shown the Figure 4.8.2, the accuracy of the repited results is really close to the original. This was done using the same samples of the previous study but with the current SVM. The use of a different SVM is the reason for the slight difference in accuracy.

Therefore it can be concluded that the developed algorithm is correct, so it is able to start testing this method.

	Radius	4 points	8 points
1.	4	90,89	92,68
	5	93,39	96,07
	6	91,79	95,54
	7	87,86	96,43
	8	91,34	96,79
2.	8	91,25	96,96

Figure 4.8: Balanced classication rate (%) using the LBP method with 4 and 8 points with different radius: 1. Previus results. 2. Repeatation of the previous results.

4.2.2 The new tests of LBP's method

Below are presented the obtained results of the new tests of the LBP method, but using the set of images introduced in Section 3.2 without the normalization of the rotation. Are also used all the tools and criteria explained in Chapter 3.

Recognition rate numbers without noise label

Firstly will be tested the efficiency of this method to recognize a number. Below is the confusion matrix in Figure 4.9. As can be seen it tends to accumulate on the diagonal the peak values.

1.	4	5	6	7	8	9	10	11	13	14	15
4	585	0	0	22	11	14	0	0	0	0	2
5	12	80	0	12	0	0	2	0	0	24	0
6	2	2	0	1	6	18	0	0	0	8	0
7	19	2	0	228	6	2	0	0	0	8	8
8	13	0	0	2	308	30	1	0	0	0	0
9	40	2	0	5	25	354	2	0	2	15	0
10	3	0	0	5	0	90	37	0	8	3	4
11	3	0	0	3	0	0	23	179	24	4	1
13	0	0	0	1	1	1	5	1	243	0	33
14	10	10	0	11	2	17	4	1	4	104	0
15	1	1	0	6	0	2	0	5	59	7	103

2.	4	5	6	7	8	9	10	11	13	14	15
4	92.27	-	-	3.47	1.74	2.21	-	-	-	-	0.32
5	9.23	61.54	-	9.23	-	-	1.54	-	-	18.46	-
6	5.41	5.41	-	2.70	16.22	48.65	-	-	-	21.62	-
7	6.96	0.73	-	83.52	2.20	0.73	-	-	-	2.93	2.93
8	3.67	-	-	0.56	87.01	8.47	0.28	-	-	-	-
9	8.99	0.45	-	1.12	5.62	79.55	0.45	-	0.45	3.37	-
10	2.00	-	-	3.33	-	60.00	24.67	-	5.33	2.00	2.67
11	1.27	-	-	1.27	-	-	9.70	75.53	10.13	1.69	0.42
13	-	-	-	0.35	0.35	0.35	1.75	0.35	85.26	-	11.58
14	11.05	5.81	-	6.40	1.16	9.88	2.33	0.58	2.33	60.47	-
15	0.54	0.54	-	3.26	-	1.09	-	2.72	32.07	3.80	55.90

Figure 4.9: LBP's confusion matrix. The first table cointains the matrix using the number of samples. The second table contains the percentages (%).

Can be observed clearly that the samples of 5 and 14 are often confused. Observing the histogram of samples, in Figure 4.10, it can be seen as more of the components (columns 1, 4, 6, 9) of their vectors do not separate cleanly the features. On the other hand, the numbers 4 and 9 are clearly distinguishable. Noting their histograms, it can be seen as many of the vector components (2, 3, 4, 8) successfully discriminate samples.

Once it is shown the performance of this method, the classification rate obtained, using only the digit labels, is:

88.69 %

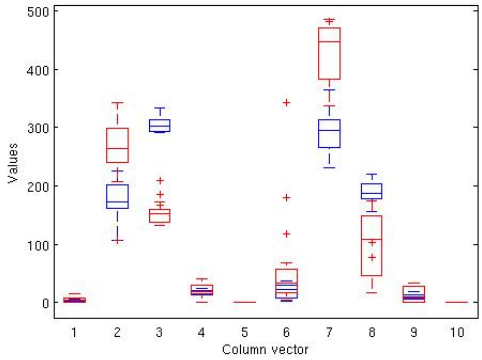

Figure 4.10: Histogram of the LPB's samples of the numbers 5 (blue) and 14 (red) with R=8 and n=8.

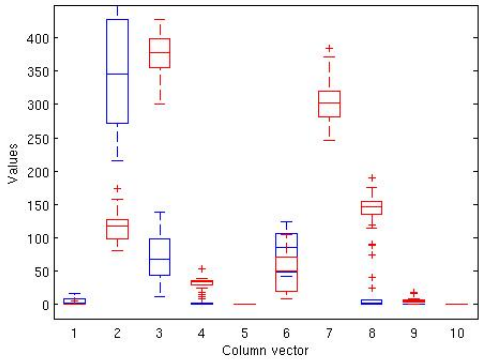

Figure 4.11: Histogram of the LPB's samples of the numbers 4 (blue) and 9 (red) with R=8 and n=8.

Recognition rate numbers with noise label

Once established that this method correctly identifies the numbers, it is necessary to prove the efficiency finding out a false detection, identifying the noise label. The classification rate is:

80.23 %

As was expected the recognition rate decreases, obtaining a value quite low to be used in a players recognition system.

Identification delay

The mean made from times needed to identify each of the samples gives the following average value of delay:

0.2885 seconds

4.2.3 LBP's conclusions

The use of the LBP based-method in the identification of numbers is remarkable. The problem is when it is required the identification of false detections of digits. This property, required in the basketball player identification system, makes the identification rate decreases significantly.

4.3 Optical character recognition

In this section, at first, are introduced the results obtained in the previous tests of the OCR method. In the absence of the images which these test had been carried out, none of these proofs have been repeated. The following results just will be compared with the new results to see if they are congruent.

4.3.1 Repetition of the OCR's previous results

The set of images that previously had tested this method contained 1284 samples. The dataset was broken down as follows:

- ▷ Team A:
  - 146 numbers "0"
  - 140 numbers "1"
  - 229 numbers "4"
  - 145 numbers "5"
  - 138 numbers "9"
- ▷ Team B:
  - 210 numbers "4"
  - 160 numbers "5"
  - 104 numbers "7"
  - 41 numbers "9"

The recognition rate obtained was 99.5%. Using the false detection label, the accuracy decreased at 98%. Unable to repeat these experiments, will be compared these results with new ones to check that are reasonable.

4.3.2 The new tests of OCR's method

Below are shown the current tests of the OCR based method. It has been used the set of images introduced in Section 3.2, with the rotation normalized. Are also used all the tools and criteria introduced in Chapter 3.

Recognition rate numbers without noise label

Firstly will be tested the efficiency of this method to recognize a number. The classification rate obtained, using only the digit labels, is:

96.65 %

The confusion matrix of this method is shown in Figure 4.12. As can be seen it tends to accumulate on the diagonal peak values.

1.	4	5	6	7	8	9	10	11	13	14	15
4	617	0	0	7	0	0	0	0	0	9	0
5	0	129	0	1	0	0	0	0	0	0	0
6	0	0	33	0	1	0	0	0	0	0	3
7	13	0	0	254	0	0	1	0	0	33	1
8	0	0	2	1	345	5	0	0	0	1	0
9	0	1	1	0	5	433	0	0	5	2	0
10	2	0	0	1	1	0	146	0	0	0	0
11	1	0	0	2	0	0	0	233	1	0	0
13	0	0	0	0	0	1	0	0	280	2	0
14	6	0	0	2	0	0	0	1	1	163	0
15	4	0	1	1	1	1	0	3	2	1	170

2.	4	5	6	7	8	9	10	11	13	14	15
4	97.47	-	-	1.11	-	-	-	-	-	1.42	-
5	-	99.23	-	0.77	-	-	-	-	-	-	-
6	-	-	89.19	-	2.70	-	-	-	-	-	8.11
7	4.30	-	-	84.11	-	-	0.33	-	-	10.93	0.33
8	-	-	0.56	0.28	97.46	1.41	-	-	-	0.28	-
9	-	0.22	0.22	-	1.12	99.11	-	-	1.12	0.45	-
10	1.33	-	-	0.67	0.67	-	97.33	-	-	-	-
11	0.42	-	-	0.84	-	-	-	98.31	0.42	-	-
13	-	-	-	-	-	0.35	-	-	98.94	0.71	-
14	3.47	-	-	1.16	-	-	-	0.58	0.58	94.22	-
15	2.17	-	0.54	0.54	0.54	0.54	-	1.63	1.09	0.54	92.39

Figure 4.12: OCR's confusion matrix. The first table cointains the matrix using the number of samples. The second table contains the percentages (%).

Recognition rate numbers with noise label

Once established that this method correctly identifies the numbers, it is necessary to prove the efficiency finding out a false detection, identifying the noise label. The classification rate is:

91.61 %

As expected the recognition rate decreases, obtaining a value too low to be used in a players recognition system.

Identification delay

The mean made from the times needed to identify each of the samples gives the following delay average value:

0.0111 seconds

4.3.3 OCR's conclusions

The results are consistent in comparison with the findings of the previous study. Using the OCR method to characterize the images used in the identification of basketball players, a sufficiently high recognition rate is gotten. Besides is also efficient recognizing false detections. Finally, the delay in the recognition is really low.

4.4 Comparative analysis

Once collected the results of each method tests, attached in Figure 4.13, can be concluded that the most efficient method is based on the OCR. It is the method that requires less time to characterize and classify an image, and making it more precisely.

	MP	LBP	OCR
Accuracy, just numbers (%)	55.51	88.69	**96.65**
Accuracy, with false detections (%)	30.01	80.23	**91.61**
Recognition delay (seconds)	3.6398	0.2885	**0.0111**

Figure 4.13: Table with the results obtained from testing the three methods.

It is important to explain the reason which all results were worse than those obtained in previous studies. The main reason is the set of binary images used. All previous studies used the same dataset, with the same fields, but in different tests. Each number had the variables shown in Figure 4.14. Basically the vector

remove obviates the samples that are clearly not a digit. The Figure 4.15 shows an example of discarded samples. Working directly with the pictures obtained from the binarization of the input RGB images, as the dataset used in this project, without removing any sample, it makes sense that the identification accuracy falls in front of a clearer dataset.

Field	Value	Min	Max
missed	1	1	1
count	421	421	421
remove	<1x421 double>	0	1
sample	<421x1024 logic...		
ref	<421x6 double>	1	7.3351e+05

Figure 4.14: Fields belonging to each number used in the previous dataset.

Figure 4.15: Display of the discarded samples of number 15 pertaining to the previous dataset.

Clearly can be seen that all methods concentrate their worst results in the numbers with fewer samples. For example, in all confusion matrix the number 6 (37 samples) is which more errors accumulate. To test this hypothesis, the 10-fold cross-validation is performed again but in those numbers with fewer samples its have been increased, trying to have an equitable number of samples for each label.

- ▷ Team A (yellow):
  - 421 numbers "4"
  - 390 numbers "5"
  - 335 numbers "8"
  - 350 numbers "9"
  - 300 numbers "10"
  - 176 numbers "11"
  - 350 numbers "14"

▷ Team B (black)

- 234 numbers "4"
- 370 numbers "6"
- 275 numbers "7"
- 46 numbers "8"
- 101 numbers "9"
- 149 numbers "11"
- 287 numbers "13"
- 368 numbers "15"

1.	4	5	6	7	8	9	10	11	13	14	15
4	624	1	0	6	0	0	0	0	0	1	1
5	0	390	0	0	0	0	0	0	0	0	0
6	0	0	36	0	1	0	0	0	0	0	0
7	12	1	0	256	0	0	0	0	0	0	3
8	0	0	2	1	346	4	0	0	0	1	0
9	0	2	2	0	6	430	0	0	5	2	0
10	0	0	0	0	0	0	300	0	0	0	0
11	1	0	0	2	1	0	0	232	1	0	0
13	0	0	0	1	0	1	0	0	280	1	0
14	2	0	0	0	0	0	0	0	0	344	0
15	3	0	1	4	0	0	0	5	0	3	352

2.	4	5	6	7	8	9	10	11	13	14	15
4	98.58	0.15	-	0.94	-	-	-	-	-	0.15	0.15
5	-	100	-	-	-	-	-	-	-	-	-
6	-	-	97.29	-	2.78	-	-	-	-	-	-
7	0.74	0.37	-	94.12	-	-	-	-	-	-	1.11
8	-	-	0.56	0.28	97.74	1.13	-	-	-	0.28	-
9	-	0.45	0.45	-	1.34	96.20	-	-	1.12	0.45	-
10	-	-	-	-	-	-	100	-	-	-	-
11	0.42	-	-	0.84	0.42	-	-	97.89	0.42	-	-
13	-	-	-	0.35	-	0.35	-	-	98.95	0.35	-
14	0.57	-	-	-	-	-	-	-	-	99.43	-
15	0.82	-	0.27	1.09	-	-	-	1.36	-	0.82	95.65

Figure 4.16: OCR's confusion matrix with an equilibrated number of samples. The first table cointains the matrix using the number of samples. The second table contains the percentages (%).

This are the obtained results using just digits, without the false detection label. The Figure 4.16 shows the new confussion matrix and the classification rate is:

97.90 %

As was expected the recognition precision is increased using more samples. This is due to a better training of SVM classifier.

Each of these methods provides an array of features extracted from an image. But what may be the result of combining the methods studied? The MP method is rejected because the delay added to the recognition is too high. So in the next section will be discussed the results combining the features of the OCR and LBP methods.

4.4.1 OCR & LBP

Using these two methods to characterize the images, it could be achieved a better accuracy than the provided just by the OCR method. Also could be achieved a greater robustness to the rotation of the images. The following tests have been performed using the dataset with the original amount of samples (the set of images introduced in Section3.2) and with the rotation unnormalized.

Considering only the numbers, without using the false detections label, can be seen in the confusion matrix (Figure 4.17) an improvement of the number classification. The classification rate obtained with the 10-fold crosvalidacion is:

96.76 %

The accuracy obtained is almost identical to the gotten using only the OCR method. This means that has been achieved a robustness to rotation. Now it will be checked the efficiency finding out a false detection, using the noise label. The classification rate is:

91.25 %

But the fact of combining the features of the two methods has a cost. An increase in the delay of identification, which is the sum of each method times:

0.2916 seconds

Therefore should be checked if this increase in the delay offsets the time required to correct the rotation of the images. As shows the Figure 4.18, in the process to prepare the input image to the statistical processing, the correction of the rotation adds less delay than that required by the combination of OCR & LBP methods to recognize the image.

1.	4	5	6	7	8	9	10	11	13	14	15
4	617	1	1	1	0	6	0	0	0	5	3
5	0	125	0	2	0	2	0	0	0	0	1
6	1	0	30	0	2	3	0	0	0	0	1
7	5	1	1	264	0	0	0	1	0	0	1
8	0	0	1	1	346	4	0	0	0	1	1
9	1	1	1	1	7	433	1	1	1	0	0
10	0	0	0	0	0	1	145	1	2	0	1
11	1	0	0	2	0	1	0	229	2	0	2
13	1	0	0	0	0	2	0	0	278	0	2
14	7	0	0	1	0	3	0	0	0	162	0
15	2	1	1	2	1	0	0	2	8	0	167

2.	4	5	6	7	8	9	10	11	13	14	15
4	97.32	0.16	0.16	0.16	-	0.95	-	-	-	0.79	0.47
5	-	96.15	-	1.54	-	1.54	-	-	-	-	0.77
6	2.70	-	81.08	-	5.41	8.11	-	-	-	-	2.70
7	1.83	0.37	0.37	96.70	-	-	-	0.37	-	-	0.37
8	-	-	0.28	0.28	97.74	1.13	-	-	-	0.28	0.28
9	0.22	0.22	0.22	0.22	1.57	96.87	0.22	0.22	0.22	-	-
10	-	-	-	-	-	0.67	96.67	0.67	1.33	-	0.67
11	0.42	-	-	0.84	-	0.42	-	96.62	0.84	-	0.84
13	0.35	-	-	-	-	0.71	-	-	98.23	-	0.71
14	4.05	-	-	0.58	-	1.73	-	-	-	93.64	-
15	1.09	0.54	0.54	1.09	0.54	-	-	1.09	4.35	0.54	90.76

Figure 4.17: OCR & LPB's confusion matrix with R=8 and n=8. The first table cointains the matrix using the number of samples. The second table contains the percentages (%).

RGB to bin with rotation normalized	RGB to bin rotation unnormalized	Time difference
0.1400	0.0441	0.0959
OCR's identification delay	**OCR & LBP's identification delay**	**Time difference**
0.0111	0.2916	0.2805

Figure 4.18: Comparison of the time differences (seconds) required to normalize the rotation of the input image and the identification by the two methods.

4.5 Conclusions

After analysing the results, it has been concluded which is the most efficient method of the three introduced. The best method is the based on OCR, having the best accuracy and that requires the less time to identify an image.

In addition it has been confirmed the hypothesis that, the more samples contains each label, the better accuracy of SVM classifier is achieved. Finally it is shown that combining the samples of OCR and LBP methods it is gotten the same accuracy as just using the OCR method, but it is required more time to recognize an image.

Chapter 5

Conclusions

The purpose of this project has been useful to compare three different procedures to identify numbers of an image. The main idea is to integrate the best method in the system that detects and recognizes players on a sport-field.

Working with the numbers of the players jerseys implies that the shapes to identify suffer rotations and non-rigid deformations. Previous studies had introduced two methods that provide features invariants to the rotation, based on matching pursuit and local binary pattern algorithm. It was necessary compare these two methods with the currently used by the Apidis project, based on the optical character recognition.

This work has succeeded in making a fair comparison of these methods, using a set of images of basketball players shirts and a common support vector machine classifier.

The method based on matching pursuit is not suitable for a number identification system. Its classification rate is really low and the time that needs to characterize the input picture is so high.

On the other hand, the method based on local binary pattern gets better results. It provides an acceptable accuracy but it has problems using a label to identify the false detections. The accuracy decreases to a rate not useful for a recognition of players system.

Finally, the method based on the opticar character recognition achieves the highest classification rate and it is also the procedure that requires less time to obtain the image features to identify the number. The only drawback of this method is that it is necessary to correct the rotation of the input images.

Combining the optical character recognition and the local binary pattern methods, the obtained features are robust to rotation. The classification rate is equal to that provided by only the optical character recognition method, but without the need to normalize the rotation of the input images. But this combination causes the identification delay increases, being the sum of each time method. Through testing, it was found that the time required to normalize the rotation is less than the increase of delay that causes the combination of the two methods.

So after doing a fair comparison with common criteria and a homogeneous environment, it has been concluded that the most efficient method to use in a player identification system is the based on the optical character recognition.

Once that has been proved the correct performance of the last stage in the players recognition and that it is working with the most efficient way, just need to keep working with the other steps to continue the development of the Apidis project.

Bibliography

[1] Adrien Dessy Bart de Spiegeleer. *Player number recognition.* PhD thesis, Universite Catholique de Louvain, 10 decembre 2009.

[2] Mairy Jean-Baptiste Bernard Guillaume, Branders Samuel. *Player Number Recognition.* PhD thesis, Universite Catholique de Louvain, 9 decembre 2009.

[3] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *http://www.csie.ntu.edu.tw/ cjlin/papers/libsvm.pdf.*

[4] Christophe De Vleeschouwer Damien Delannay, Nicolas Danhier. Detection and recognition of sports(wo)men from multiple views. *IEEE International Conference on Distributed Smart Cameras (ICDSC 2009)*, Septembre 2009.

[5] Nicolas Danhier. *Detection et reconnaissance de joueurs de basket-ball a l'aide d'un reseau de cameras.* PhD thesis, Universite Catholique de Louvain, 2009.

[6] Collignon Thomas Dotremont Olivier. *Image processing and computer vision.* Ecole Polytechnique de Louvain, 22 MARS - Q1 : 2009 - 2010.

[7] Qingming Huang Wen Gao Guangyu Zhu, Changsheng Xu. *Automatic multi-player detection and tracking in broadcast sports video using support vector machine and particle filter.* Harbin Institute of Technology, Harbin, P.R. China, 2006.

[8] T. Taxt O. D. Trier, A. K. Jain. *Feature extraction methods for character recognition.* A survey, Pattern Recognition, vol. 29, pp. 641-662., 1996.

[9] Shuqiang Jiang Yang Liu Wen Gao Qixiang Ye, Qingming Huang. *Jersey number detection in sports video for athlete identification.* PhD thesis, Institute of Computing Technology of Chinese Academy of Sciences, 2005.

[10] Pierre Vandergheynst Rosa M. Figueras i Ventura. *Low-Rate and Flexible Image Coding With Redundant Representations.* IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 15, NO. 3, MARCH 2006.

[11] Hideo Saito Sachiko Iwase. *Tracking Soccer Players Based on Homography among Multiple Views.* PhD thesis, Department of Information and Computer Science, Keio University, Yokohama, Japan, 2003.

[12] Abdenour Hadid Timo Ahonen and Matti Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(12) :2037-2041.*, 2006.