Face Recognition Using SIFT and Binary PSO Descriptors

Lanzarini Laura, La Battaglia Juan, Maulini Juan, Hasperué Waldo
III- LIDI (Institute of Research in Computer Science III-LIDI)
School of Computer Science – UNLP - La Plata, Bs.As. - Argentina
{laural, juanlb, jmaulini, whasperue}@lidi.info.unlp.edu.ar

Abstract. *In this paper, a strategy for face recognition based on SIFT descriptors of the images involved is presented.*
In order to reduce the number of false positives and computation time, a selection of the most representative feature descriptors is carried out by applying a variation of the binary PSO method.
The results obtained show that the strategy proposed is better than the direct application of SIFT descriptors.

Keywords. Face recognition; SIFT Features; Swarm Intelligence

1. Introduction

Face recognition is a biometric technique that is widely used in various areas such as security and access control, forensic medicine, and police controls. It involves determining if the image of the face of any given person matches any of the face images stored in a database. This problem is hard to solve automatically due to the changes that various factors, such as facial expression, aging and even lighting, can cause on the image.

In this paper, a method using only those SIFT descriptors that best represent the image is proposed. Good recognition results are achieved while solving the two major problems of this characterization method: false positive detection and the time required for the recognition process.

The selection of SIFT descriptors is carried out by means of a variation of binary PSO (Particle Swarm Optimization), and it is applied only to database image descriptors. Therefore, SIFT descriptors processing is done before the recognition stage of the process.

This paper is organized as follows: In Section 2, a brief description of previous related works using similar techniques is included; in Section 3, the method that allows obtaining SIFT descriptors from an image is described; whereas in Section 4 some clarifications regarding the binary PSO variation used are presented. In Section 5, implementation details are provided, and in Section 6 the results obtained are described. Finally, in Section 7 the conclusions obtained are presented.

2. Related work

There are currently various solutions to this problem that use SIFT descriptors.

It has been shown [1] that using SIFT descriptors for the face recognition process is better than Eigenfaces and Fisherfaces algorithms. Training datasets were of various sizes, which allowed establishing that performance decreases as dataset size decreases. As regards the significant number of SIFT descriptors required for a reliable comparison, it was observed that, with a lower number of descriptors, performance is better than that obtained with Eigenfaces and Fisherfaces.

In order to tackle the issue of comparing very long feature vectors for all images in a database, a biased classification of the features that make SIFT descriptors, is proposed and used to reduce the length of SIFT descriptors used for face recognition [6]. Thus, the number of comparisons is reduced and the recognition process is faster. This process also filters out those descriptors that are irrelevant for face recognition, thus increasing recognition accuracy.

On the other hand, a face recognition algorithm that uses the binary PSO algorithm to explore the solution space for an optimum subset of features in order to increase recognition rate and class separation is presented in [8]. This algorithm is applied to feature vectors extracted using the Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT).

3. SIFT Features

In [5], Lowe defined a method to extract features from an image and use them to find

Proceedings of the *ITI 2010 32nd Int. Conf. on Information Technology Interfaces,* June 21-24, 2010, Cavtat, Croatia

matches between two different views of the same object. These features, called SIFT (Scale Invariant Feature Transform) features, are invariant to image scale and rotation, and quite invariant to affine distortion, as well as changes in point of view and lighting. They are also highly distinctive.

The process to determine SIFT features for an image consists in four steps:

- First, the location of potential points of interest within the image is determined. These points of interest correspond to the extreme points calculated from plane subsets of Difference of Gaussian (DoG) filters applied to the image at different scales.
- Then, the points of interest whose contrast is low are discarded. This is an improvement from the definition in [4].
- After this, the orientation of relevant points of interest is calculated.
- Using the previous orientations, the environment is analyzed for each point and the corresponding feature vector is determined.

As a result of this process, a set of 128-length feature vectors is obtained. These feature vectors can then be compared with those from another image of the same object with a different scale, orientation, and/or point of view.

This comparison can be done directly by measuring the distance and establishing a similarity threshold.

More detailed information about this method is available in [5].

4. Binary PSO

Particle Swarm Optimization (PSO) was originally defined by Eberhart and Kennedy in 1995 [2]. It is a populational search algorithm based on the simulation of the social behavior of some species, such as birds and fish. Each element of the population is referred to as a particle, and is represented by three vectors and two real values: One of the vectors indicates the position of the particle within the search space, the second vector (also called velocity vector) represents the last position change, and the third one is a copy of the best position of the particle found so far. The real values associated to the particle are a suitability value and the value corresponding to the best solution found so far.

Each particle moves within the solution space taking into account its current speed, the best solution found so far, and the position of the best individual of the entire population (gbest PSO version).

However, PSO was originally developed for a space of continuous values and it therefore poses several problems for spaces of discrete values where the variable domain is finite. Kennedy and Eberhart [3] presented a discrete binary version of PSO for these discrete optimization problems. In binary PSO, each particle represents its position with binary values (0 or 1). Speed is defined as the probability that a particle has of changing its value to 1.

In novel binary PSO [7], the speed of a particle is no longer defined as the probability of changing to 1, but as the probability of changing from the previous state to its complementary value. In this new definition, both the speed of a particle and its parameters have the same role as in the continuous version of PSO. The best position visited by particle (P_{ibest}) and the best global position (P_{gbest}) are updated as in the continuous or binary version of PSO.

However, two vectors are introduced for each particle: V_i^0 is the probability that bits of the particle change to zero, and V_i^1 is the probability that they change to one. Since an inertia term is used in the equation that updates these speeds, they are not complementary. The probability of a change occurring in the j^th^ bit of the i^th^ particle is defined as follows:

$$V_{ij} = \begin{cases} V_{ij}^1 & if\ x_{ij} = 1 \\ V_{ij}^0 & if\ x_{ij} = 0 \end{cases}$$

x_{ij} being the j^{th} bit in the position vector of the i^{th} particle.

These vectors are updated in the following way: If the j^{th} bit at the best global position is 0 (zero), or if the j^{th} bit at the best solution found for the particle is zero, the speed (V_{ij}^0) is increased and the probability of changing to one decreases. On the contrary, if the j^{th} bit at the best global position is 1, or if the j^{th} bit at the best solution found for the particle is 1, V_{ij}^1 is increased and V_{ij}^0 decreases. In this approach, change direction to 1 or 0 for each bit is considered separately.

After updating the speed of the particles, the speed of change is obtained.

A normalization process is also carried out by applying the sigmoid function to move each particle to a new position.

More detailed information about this method is available in [7].

5. Proposed method

In order to perform face recognition, the method proposed uses a minimum-size database formed by the subset of most representative SIFT descriptors. Thus, the computing time required to make the necessary comparisons and detection of false positives is reduced. This selection process is performed before the recognition process; therefore, it does not affect the response time for the end user. The selection process is described in Section 5.1.

The recognition of a new face involves the following steps:

- Calculating the SIFT descriptors corresponding to the input image.
- Comparing each descriptor in the database with the set of descriptors corresponding to the new face. Matches are accumulated not by image but rather by the number of the person to whom the database descriptor corresponds.
- The new face will correspond to the person with the highest number of accumulated matches.

It should be noted that the comparison of each database descriptor with the set of descriptors corresponding to the image to be recognized is a purely parallel task. If a parallel computation architecture were available, the database of SIFT descriptors could be partitioned so that each processor would have the information corresponding to one person, or, even better, to one image. Thus, the calculation of the number of matches found would be faster. As regards the recognition of the new face, a minimum threshold of matches can be used to identify faces that have no matches in the database.

In the following section, the process of selecting the SIFT descriptors that will be included in the database is described.

5.1. Building the database

The method begins by obtaining all SIFT descriptors corresponding to each input image.

The selection of the most representative SIFT descriptors is carried out by applying a variation of the method defined in [7], based on subpopulations of particles. In this case, the number of populations to use matches the number of images in the database. The length of the position vector for each particle of a population is determined by the number of SIFT descriptors of the corresponding image. Therefore, the length of particles from different populations can be different.

That is, the vector of the j^{th} particle in subpopulation i has the following form

$$X^i_j = (x^i_{j1}, x^i_{j2}, \ldots . x^i_{jmi}) \qquad (**)$$

where

- m_i is the number of SIFT descriptors of image i.
- x^i_{jk} is 1 if the k^{th} SIFT descriptor must be included in the database and 0 if it is not.

This speciation criterion allows calculating the movement of each particle using only the SIFT descriptors from one image. Thus, each population searches a different part of the solution space.

The final solution is obtained by concatenation of the best individuals of each population. This can be expressed as follows

$$X = (X^1_{best}, X^2_{best}, \ldots X^M_{best},)$$

where M is the number of different images used to form the database and X^i_{best} is the best individual in the i^{th} subpopulation.

With respect to the usual parameters of PSO:

- In each iteration, the value of w decreases, as mentioned in [3].
- Elitism was used so that, if moving individuals does not allow at least maintaining the highest fitness value found thus far, the best individual of the previous iteration regains its previous position and the fitness value lost.

The algorithm terminates when the maximum number of iterations was reached or when, after a certain number of consecutive iterations, the best fitness value has not changed.

This is summarized by the pseudo-code shown in Figure 1.

5.2. Assessing the fitness value of each particle

In this section, the method used to measure the fitness value for each particle is described.

An expression that helps reducing the number of false positives must be used. Therefore, its value increases when the selected vector has a match in an image of the corresponding subject, and it decreases when there are no matches.

```
ImList ← {list of input images}
W_ini ← initial acceleration
W_fin ← final acceleration
MAX_ITERA ← maximum number of iterations
Popsize ← size of each sub-population
Pop = CreateInitialPops(PopSize, ImList);
ite=0;
while {termination condition is not reached}
  ite = ite + 1
  w = w_ini – (w_ini-w_fin)*ite/MAX_ITERA
 for i=1,Number_of_Input_Images
    %--- each image has its own population ----
    Save the best particle of subpopulation Pop(i)
    Move the particles in Pop(i) based on Binary PSO
    Assess fitness for all particles in Pop(i)
    Apply elitism restoring the best particle if needed
 end for
 Output ← {concatenation of the best particles in
             each subpopulation}
 Output_Fitness ← {sum of fitness values corresponding
                    to the best particles of each
                    subpopulation}
end while
```

Figure 1. Pseudo-code of the method proposed to select the SIFT descriptors that will form the database.

Be X^i_j the position vector of the j[th] particle of sub-population i, defined in (**)

Be $C1^i_{jk}$ the total number of matches between the k[th] SIFT descriptor of image i and the rest of the images that correspond to the subject represented by image i.

Be $C2^i_{jk}$ the total number of matches between the k[th] SIFT descriptor of image i and the images that correspond to subjects other than that represented by image i.

The fitness value of the j[th] particle of sub-population i is calculated as

$$Fit^i_j = \sum_{k=1}^{m} x^i_{jk} * \left(\alpha_1 * C1^i_{jk} - \alpha_2 * C2^i_{jk}\right)$$

where α_1 and α_2 are constants with values between (0,1) and represent the significance of each term within the expression. As mentioned above, x^i_{jk} is 1 if the k[th] SIFT descriptor must be included in the data base, and 0 if it is not.

6. Results Obtained

Table 1. Results for the YALE database. The columns labeled (1) indicate the percentage of correct matches. The columns labeled with (2) show the average number of SIFT descriptors used per input image.

%	Proposed method		SIFT standard	
	(1)	(2)	(1)	(2)
10%	58.87%	46.71	58.72%	73.07
20%	80.01%	43.21	80.48%	73.56
30%	87.90%	39.99	88.29%	73.31
40%	89.83%	40.49	90.21%	71.84
50%	89.64%	39.49	90.12%	73.25
60%	90.91%	38.46	89.39%	72.99
70%	95.76%	37.37	94.12%	72.67
80%	96.12%	36.81	96.67%	72.62
90%	98.10%	36.25	97.14%	72.59
99%	100.00%	36.22	100.00%	72.63

Measurements were carried out using two databases obtained from [9]. The first of these is the YALE faces database, containing 165 images of 15 different subjects (11 images per person). Each image has a resolution of 320x243 pixels.

The second database used was the AT&T faces database, containing 400 images of 40 people (10 images per individual). The size of each image is 112x92 pixels.

The available images were divided in two parts:

– <u>Subset of input images</u>, whose descriptors will be selected by applying the method proposed in Section 5.
– <u>Subset of test images</u> that will be compared with the selected SIFT descriptors for recognition.

The initial SIFT descriptors for each image were determined with a threshold of 0.5, as recommended in [4].

In both cases, the parameters used by PSO were the following:

– Initial and final inertia values: 1.2 and 0.2, respectively.
– Maximum number of iterations = 500,
– α_1= 1/(number of input images),
– α_2= 16/(number of input images),

Figure 2 shows the original SIFT descriptors in the top row of the images and the descriptors selected by the proposed algorithm in the bottom row.

Table 1 shows the results obtained with the YALE faces database using different percentages of input images. The values represent the average results of 30 independent executions for each percentage. Before calculating the SIFT descriptors, the image size was reduced to 160x121.

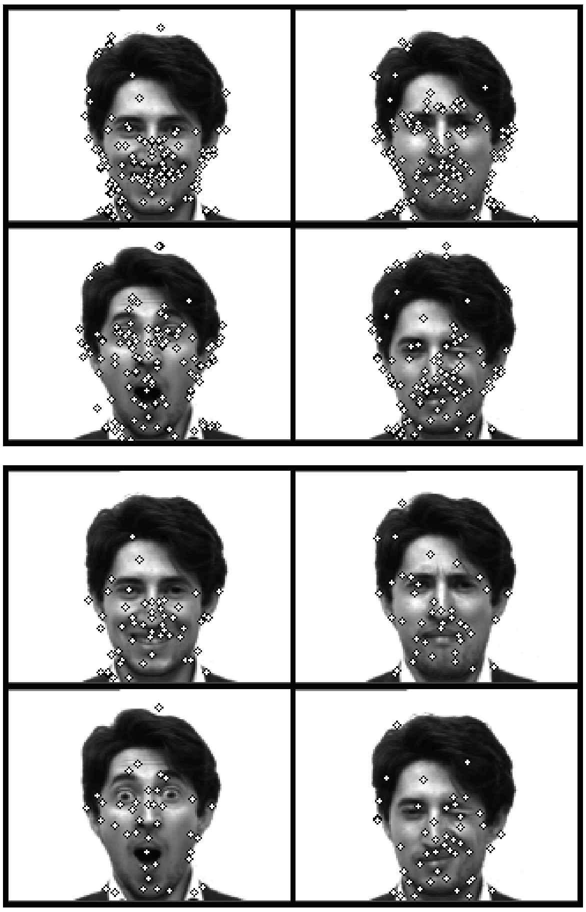

Figure 2. SIFT descriptors of a person of the YALE database. The top row shows all descriptors found while the bottom row shows only the descriptors selected by the proposed method.

As it was to be expected, the percentage of correct matches in the classification decreases as the number of input images decreases. However, the average number of SIFT descriptors used in each case is much lower than when using all the descriptors generated by Lowe's method [5] with an equivalent rate of correct matches.

As regards identification results, Figure 3 shows that the method proposed (indicated in the figure as SIFT+PSO), when applied to the YALE and AT&T databases, has a greater rate of correct matches than the conventional SIFT method. In the figure, the results obtained from 30 independent runs using 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 99% of the images included in the databases were used for training, and the corresponding remaining images for testing.

The improvement introduced by the method could not be achieved without a reduction in the number of false positives, since recognition is based on selecting the person with the highest number of matches. The greater the difference between the two best options, the more conclusive the response obtained is. In Figure 4, the average difference between the option with the highest number of matches and the second best option (considering the total number of matches) is shown. For instance, the value of 89.12% obtained when using the method proposed over 90% of the images in the YALE database indicates that, in average, the largest amount of matches found for an image is almost 90% apart from the second highest value of the total number of matches. That is, these are very distant values, meaning that recognition is conclusive. It is also a much higher value than the 73% obtained with the conventional SIFT method. This behavior is also observed with the AT&T database, also represented in Figure 4.

It can also be observed that this degree of certainty decreases as the number of input images is decreased.

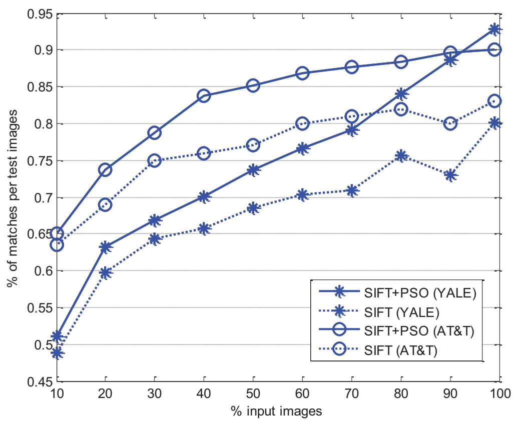

Figure 3. Percentage of matches for test images using the method proposed (SIFT+PSO) and the original SIFT method for various percentages of images from the YALE and AT&T databases

7. Conclusions

A face recognition mechanism based on SIFT features that allows reducing the size of the database by using a variation of binary PSO has been described.

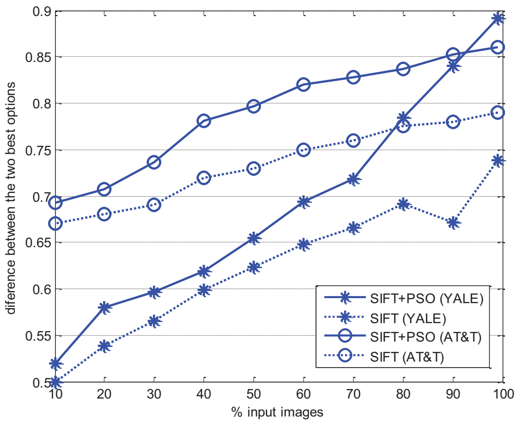

Figure 4. Average value per image of the difference between the two highest values of correct matches, divided by the total number of matches found for the YALE and AT&T databases.

The tests carried out with the YALE and AT&T databases have allowed reaching considerable reduction rates – 50% in both cases. The results obtained are equivalent to those that can be achieved using the entire SIFT features database, but only half the comparisons are required due to the selection carried out by means of a variation of binary PSO.

The parameters involved still need to be thoroughly analyzed in order to determine if a more precise adjustment would allow reducing the maximum number of iterations needed to reach an optimum selection of descriptors.

The parallelization of the solution proposed also poses an interesting analysis.

8. References

[1] Aly Mohamed. Face Recognition using SIFT Features. CNS186 Term Project Winter 2006.

[2] J. Kennedy, and R. Eberhart, “Particle Swarm Optimization”, IEEE International Conference on Neural Networks (Perth, Australia), IEEE Service Center, Piscataway, NJ, IV, pp. 1942-1948, 1995.

[3] Kennedy, J.; Eberhart, R.C. “A discrete binary version of the particle swarm algorithm”, IEEE International Conference on Systems, Man, and Cybernetics, 1997.

[4] Lowe, D.G. 1999. Object recognition from local scale-invariant features. In International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157

[5] David G. Lowe. Distinctive image features from scale-invariant keypoints. International. journal of computer vision, 60, 2004

[6] A. Majumdar and R. K. Ward, DISCRIMINATIVE SIFT FEATURES FOR FACE RECOGNITION, anadian Conference on Electrical and Computer Engineering, 2009. CCECE’09. Pp 27-30. ISSN 0840-7789.

[7] Mojtaba Ahmadieh Khanesar. A novel Binary Particle Swarm Optimization. 18th Mediterranean Conference on Control & Automation. 2007. T33-001.

[8]Rabab M. Ramadan and Rehab F. Abdel – Kader. Face Recognition Using Particle Swarm Optimization-Based Selected Features. International Journal of Signal Processing, Image Processing and Pattern Recognition. Vol2, nro. 2, June 2009.pp51-65.

[9] http://www.face-rec.org/databases