

Indexing Of Three Dimensions Objects Using GIST, Zernike & PCA Descriptors

Driss Naji^{*}, M. Fakir^{}, O. Bencharef^{***}, B. Bouikhalene^{****}**

^{*} Faculty of Science and Technology /Informatics, BeniMellal, Morocco. Email: naji.driss@gmail.com

^{**}Faculty of Science and Technology /Informatics, BeniMellal, Morocco. Email: fakfad@yahoo.fr

^{***}Faculty of Science and Technology /Informatics, BeniMellal, Morocco. Email: bencharef98@gmail.com

^{****}Faculty of Science and Technology /Informatics, BeniMellal, Morocco. Email: bbouikhalene@yahoo.fr

Article Info

Article history:

Received Jul 22, 2012

Revised Nov 15, 2012

Accepted Jan 05, 2013

Keyword:

Gist descriptor

Zernike moments

PCA

hybrid descriptors

SVM

Neural Networks (NN)

ABSTRACT

The fact of using the classic descriptors such as Zernike Moment and Gist for a large data base has never been a satisfying method for perfect recognition rates. In this paper, we came up with a different approach based on the combination of the different descriptors already mentioned, it is the result of a comparative study of the different descriptors and the different combinations (Zernike + Gist, Zernike + PCA, Gist + PCA) in terms of recognition rate. Eventually, we have deduced that the combination of Zernike moment with Gist descriptors ended up to be the best hybrid description. For the recognition process, we opted for support vector machine (SVM) and Neural Networks (NN). We illustrate the proposed method on 3D objects using a 2D/3D approach based on characteristic views.

*Copyright © 2013 Institute of Advanced Engineering and Science.
All rights reserved.*

Corresponding Author:

Mohamed FAKIR,
Departement of Electrical and Computer Engineering,
Faculty of Sciences and Technics,
University Sultan Moulay Slimane.
Email: fakfad@yahoo.fr

1. INTRODUCTION

A central problem in object recognition resides in the complexity of real objects (a complex shape representation and a non-uniform color distribution) [1 2 3 4 5]. Without forgetting the impact of changing the angle of view .A recognition system is composed of:

- Extraction of the characteristics.
- Learning (indexing).
- Recognition (Research) [6 7].

The feature extraction is a crucial step in the recognition process. In this paper we tried to verify the complementarity of three categories of descriptors and we used the literature to choose the best descriptor of each category :

- For descriptors based on color description, we chose the Gist descriptors
- For invariant moments descriptors we vote for Zernike moments
- For data-analays descriptors we chose the Principal component analysis (PCA)

We propose a hybrid approach based on the combination of descriptors, including Zernike descriptor + Gist, Gist + PCA and PCA + Zernike.

To test our approach, we opted for two classifiers well-known for their robustness: Neural Networks and Wide Margin Separators (SVM).

2. GIST DESCRIPTOR

GIST descriptor, constructed by (Oliva and Torralba, 2001) [8 9] for a comprehensive description of the scene. In its original description (and in the implementation we use), it is close enough to the Gabor filter bank. The use of GIST occurs in three stages:

- Calculation of a descriptor "raw" high-dimensionality of learning some basic
- Calculation of a PCA on the training set
- For a new image, the descriptor is calculated gross, and it is projected with the PCA to reduce dimensionality.

The descriptor "gross" is constructed as follows:

- We move the image in a Gabor filter bank with N_σ scales N_θ orientations scale we obtain $N = N_\theta \times N_\sigma$ pictures
- Each image is cut to $M \times M_{sub}$ images
- calculating the energy of each sub-picture; is obtained thus a vector of size $M \times N$

For color images, it does so for each channel, the descriptor size will be $3 \times M \times N$.

3. THE PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA is a method of family data analysis and more generally of multivariate statistical [10], which is to transform these variables (called "correlated" in statistics) in new variables decorates each other. These new variables are called "principal components", or axes. It allows the practitioner to reduce the information in a more limited number of components in the original number of variables.

This is an approach that is both geometric (representation of variables in a new geometric space along directions of maximum inertia) and statistics (search for independent axes explaining the most variability - variance - data). The algorithm can be summarized as follows:

- ✓ This explode matrix eigenvectors, eigenvalues $\{p_i, \lambda_i\}$
- ✓ Center of data
- ✓ To order the eigenvalues in descending order
- ✓ Construct the covariance matrix Σ
- ✓ The subspace of dimension q that best represents the data in the sense of mean square error is generated by the matrix:

$$P = (p_1, p_2, \dots, p_{n-1}, p_q)$$

Where $\{p_1, p_2, \dots, p_{n-1}, p_q\}$ are the eigenvectors associated to the q largest eigenvalues.

- ✓ All the main components are written in matrix form:

$$C = XP = (c_1, c_2, \dots, c_{n-1}, c_q)$$

4. ZERNIKE MOMENTS

Zernike moments were introduced by F. Zernike in 1934 [11]. In the field of information processing, Zernike moments have been used extensively for their orthogonality property that allows the generation of non-redundant descriptors and their invariance properties in translation, rotation and scale. Thus, we find the Zernike moments in many works on recognition of images of people [12], image indexing in the database, analysis and description of the form of 2D or 3D object.

The formulation of these moments is given by:

$$Z_{nm} = \frac{n+1}{\pi} \iint_{xy} f(x, y) \cdot [V_{nm}(x, y)]^* dx dy \quad (1.1)$$

[...]* is used to indicate the complex conjugate value, here n is the order of decomposition ($n=0, 1, 2, \dots, \infty$), also known as radial order and m the number of repetitions of the decomposition or azimuthal frequency for a given order n .

The order and repetition are bound by the following two conditions:

$$n-|m| \text{ always even and } |m| \leq n \quad (1.2)$$

$V_{nm}(x, y)$ represents the Zernike polynomials constituents based orthogonal projection. They are written in general polar representation in the following form:

$$V_{nm}(r, \theta) = \mathcal{R}_{nm}(r)e^{-im\theta} \quad (1.3)$$

Or $\mathcal{R}_{nm}(r)$ are polynomials radial of the form:

$$\mathcal{R}_{nm}(r) = \sum_{k=|m|}^n r^k \frac{(-1)^{(n-k)/2} (n+k)!}{\left(\frac{n-k}{2}\right)! \left(\frac{m+k}{2}\right)! \left(\frac{k-m}{2}\right)!} \quad (1.4)$$

The application of Zernike moments to a discrete function $h(x, y)$ (such as for example) requires rewriting (1.1) as follows:

$$Z_{nm} = \frac{n+1}{\pi} \sum_x \sum_y h(x, y) [V_{nm}(x, y)]^* \quad (1.5)$$

Where $x^2 + y^2 \leq 1$

5. RESEARCH METHOD

We apply the classical phases of learning and decision as follows:

- The image is resized to 128x128 pixels.
- The Gist descriptor is calculated for color images.
- The Zernike moments are calculated from the image for each color channel.
- The set of descriptors (512 values for descriptors and GIST 21 pour Zernike moments) and 533 for the combination of both PCA and (MZ) and 10 with PCA (GIST) with 21 used to power the classifier.
- The classifier used is a neural network with an intermediate layer of 50 neurons fixed loan for all tests.
- SVM[13] was used as the second classification method for comparison with the results obtained with the neural network.

6. RESULTS AND ANALYSIS

6.1. Performance Evaluation

We present in this part the results obtained from the descriptors described above, applied to the Coil100 database, Performance has been systematically evaluated by the methodology described in section 5.

6.2. COIL-100 database

The database COIL [14] (Figure 1) contain 100 color images of objects. For each object, there are 72 images of rotating object (5° between each image).



Figure 1. Different objects in the database COIL-100

6.3. Analysis of results

Indexing and research methods must have the ability to give results even with big data base. For this, we presented the classification results in a comparison table (Table 1) between the hybrid descriptors and each one of Zernike, Gist & PCA descriptors.

Table 1. Recognition rates obtained by using SVM on the basis COIL for each family of descriptors.

Image	M.Zernike	PCA (Ze,14)	D. Gist	PCA (Gist, 100)	Gist+ Zernike	PCA
400	98,75%	93%	99,25%	98,75%	99%	90%
800	92,13%	24,25%	94,25%	97,88%	99,12%	80,40%
1200	92%	18,37%	99,50%	93,42%	99,25%	79,01%
1600	69,31%	3,13%	86,81%	84,50%	91,56%	68,12%

The best recognition rate is obtained for a combination of Gist descriptors and Zernike moments with an error of 1% for a database of 400 images. All combination gives good results for small database, for large database, the positive impact of hybrid methods is clear.

Also we can notice that Gist descriptors are more robust than Zernike moments when it comes to color objects.

When applying PCA on the descriptors, we find the following two properties:

- ✓ A reduction in computing time and recognition rate (PCA and Zernike).
- ✓ A reduction in computing time and stability but a recognition rate (PCA and Gist).

We also studied the influence of the number of images used during the learning phase. Indeed, in real applications, it is common to have only a very small number of images per object. It is therefore important that the method is sufficiently efficient in this case. Figure 2 shows the resulting error. With only 45% of training images, Gist descriptors have an error of only 3% while the error is 4% higher with the Zernike moments. The best convergence is obtained with a combination of both approaches the error is less than 1%.

Tests were made by incorporating a binary image noise (pepper and salt). Hybrid descriptors can tolerate up to 30% without substantially disturbing noise ratios.

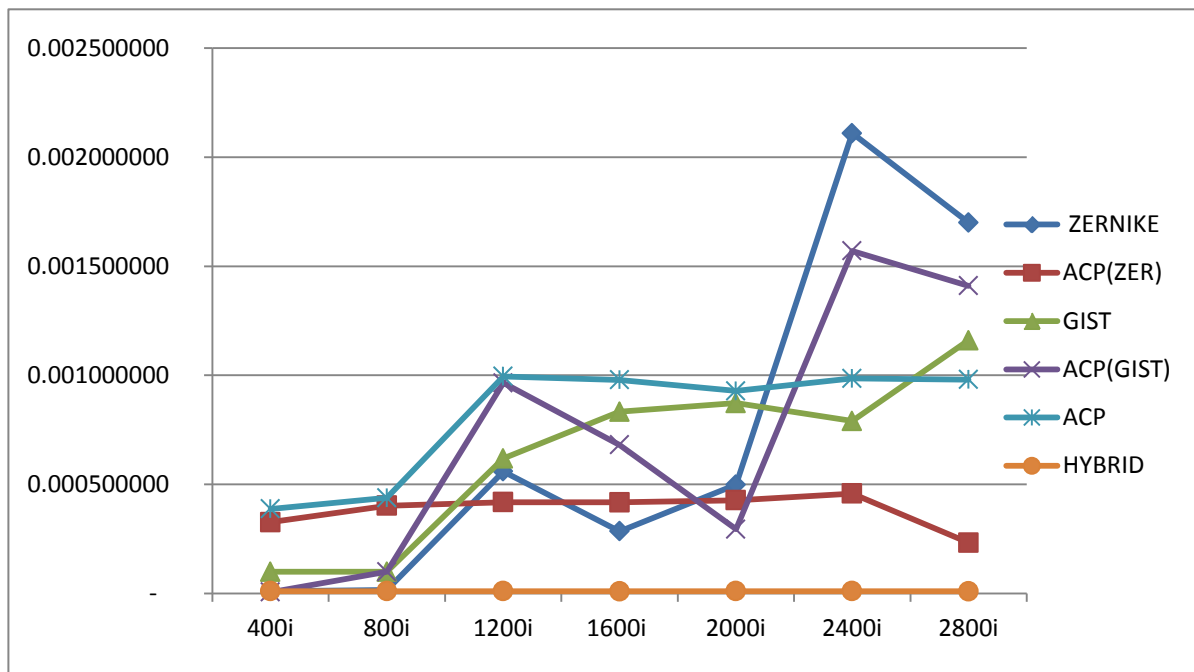


Figure 2. Classification error obtained by using NN on the basis COIL for each family of descriptors

We proposed and compared two approaches to global color object recognition: neural networks like multilayer perceptron (NN) and Support vector machines (SVM). The NN is already well known for this kind of work, and the SVM are currently emerging in this area for ease of use and excellent results. The comparison with SVM or giving advantage, and the use of SVM as a classifier makes the results more accurate, Table 2 shows the result obtained by SVM Compared to the results obtained by NN.

Table 2. Recognition rate for SVM and with NN (1200 images)

	NN	SVM
Zernike	92%	97,65%
PCA (Zer,14)	18,37%	67,33%
Gist	99,50%	100,00%
PCA(Gist,100)	93,42%	96,49%
Gist+Zernike	99,25%	100,00%
PCA	79,01%	82,60%





7. CONCLUSION

In this work, we used three extraction methods and we evaluated their performance hybrid. The results of this evaluation showed that the hybrid method (GIST Zernike +) marks the best recognition rate compared to other descriptors. The assessment of the results was carried out by neural networks and SVM

REFERENCES

- [1] Mokhtarian F, Abbasi, S et Kittler J. *Robust and Efficient Shape Indexing through Curvature Scale Space*. Dans Proceedings of the sixth British Machine Vision Conference (BMVC '96). Edinburgh, Royaume-Unis. 1996: 53-62.
- [2] J Koenderink and A van Doorn. The internal representation of solid shape with respect to vision. *Biol. Cybern.*, 1979.
- [3] Costa, MS et Shapiro, LG. 3D Object Recognition and Pose with relational indexing. *Computer Vision and Image Understanding*. 2000; 79: 364-407.
- [4] M Petrou and A Kadir. "Affine invariant features from the trace transform". *IEEE Transaction on Pattern Analysis and Machine Intelligence*. 2004; 26(1): 30-44.
- [5] F Ghorbel. "A complete invariant description for gray-level images by the harmonic analysis approach". *Pattern Recognition Letters*. 1994; 15: 1043-1051.
- [6] MJ Tarr and HH Bulthoff. Is human object recognition better described by geon structural descriptions or by multiple views? comment on biederman and gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance*. 1995.
- [7] Bluthoff HH and Edelman S. *Psychophysical support for a two-dimensional view interpolation theory of object recognition*. Proc. Natl. Acad. Sci. USA. 1992; 89: 60-64.
- [8] A Torralba, KP Murphy, and WT Freeman. Sharing visual features for multiclass and multiview object detection. *TPAMI*. 2007
- [9] Oliva A and Torralba A. Modeling the shape of the scene: a holistic representation of the spatial envelope. *International Journal of Computer Vision*. 2004; 42(3): 145-175.
- [10] J Hérault, C Jutten, and B Ans. "Détection de grandeurs primitives dans un message composite par une architecture de calcul
- [11] F Zernike. "Diffraction theory of the cut procedure and its improved form, the phase contrast method". *Physica*. 1934; 1: 689-704.
- [12] neuromimétique en apprentissage non supervisé," in Actes du Xe colloque GRETSI, Nice, France. 1985 ; 2 : 1017-1022.
- [13] BE Boser, IM Guyon, V Vapnik. "A Training Algorithm for Optimal Margin classifiers". proc. Fifth Ann. Workshop Computational Learning theory, ACM, Press. 1992: 144-152.
- [14] <http://www.cs.columbia.edu/CAVE/>

BIOGRAPHIES OF AUTHORS

	<p>Driss Naji, obtained he's Master degree in computer science in 2012 from the faculty of Science and Technology, Beni Mellal. Currently he is a PhD student at the Center of Doctoral Studies in the Faculty of Science and technology of Beni Mellal. His research interest include image processing, pattern recognition and artificiel intelligence.</p>
	<p>Pr Mohamed Fakir: obtained a degree in Master of Electrical Engineering from Nagaoka University of Technology in 1991 and a Ph.D. degree in electrical engineering from the University of Cadi Ayyad, Morocco. He was a team member in Hitachi Ltd., Japan between 1991 and 1994. He is currently a professor at the Faculty of Science and Technology, University Sultan Moulay Slimane, Morocco. His research interest include image processing, pattern recognition and artificiel intelligence.</p>
	<p>Omar BENCHAREF. Obtained the DESS degree in 2007 from the University of Cadi Ayyad Marrakech Morocco. Currently he is a PhD student at the Center of Doctoral Studies in the Faculty of Science and technology of Beni Mellal. His research concerns image processing & pattern Recognition.</p>
	<p>Belaid BOUIKHALEN obtained a Ph.D. degree in Mathematics in 2001 and a degree of Master in Computer science in 2005 from the University of Ibn Tofel Kenitra, Morocco. He is currently a professor at University Sultan Moulay Slimane, Morocco, His research focuses on mathematics and applications, decision information systems, e-learnig, pattern recognition and artificial intelligence.</p>