

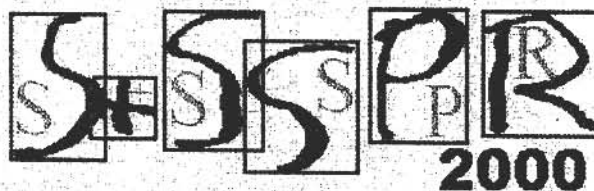
Lecture Notes in Computer Science

1876

Francesc J. Ferri José M. Iñesta
Adnan Amin Pavel Pudil (Eds.)

Advances in Pattern Recognition

Joint IAPR International Workshops
SSPR 2000 and SPR 2000
Alicante, Spain, August/September 2000
Proceedings



Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Francesc J. Ferri
University of València, Department of Computer Science
46100 Burjassot (València), Spain
E-mail: ferri@uv.es

José M. Iñesta
University of Alicante, Department of Computer Languages and Systems
03071 Alicante, Spain
E-mail: inesta@dlsi.ua.es

Adnan Amin
University of New South Wales, School of Computer Science and Engineering
Sydney, NSW 2052, Australia
E-mail: amin@cse.unsw.edu.au

Pavel Pudil
Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation
182 08 Prague 8, Czech Republic
E-mail: pudil@utia.cas.cz

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme
Advances in pattern recognition : proceedings : joint IAPR
international workshops / SSPR 2000 and SPR 2000, Alicante, Spain,
August 30 - September 1, 2000. Francesc J. Ferri ... (ed.). - Berlin ;
Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ;
Paris ; Singapore ; Tokyo : Springer, 2000
(Lecture notes in computer science ; Vol. 1876)
ISBN 3-540-67946-4

CR Subject Classification (1998): I5, I4, I2.10, I.2

ISSN 0302-9743

ISBN 3-540-67946-4 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH
© Springer-Verlag Berlin Heidelberg 2000
Printed in Germany

Typesetting: Camera-ready by author, data conversion by PTP-Berlin, Stefan Sossna
Printed on acid-free paper SPIN 10722379 06/3142 5 4 3 2 1 0

Preface

This volume contains all the papers presented at SSPR 2000 and SPR 2000, held at the University of Alicante, Spain, August 30 – September 1, 2000.

This was the second time these two technical workshops were held back-to-back. SSPR 2000 was the eighth meeting of the international workshop on Structural and Syntactic Pattern Recognition and SPR 2000 was the third international workshop on Statistical Techniques in Pattern Recognition.

These workshops have been traditionally sponsored by two of the most representative technical committees of the International Association of Pattern Recognition (IAPR): the TC2 and TC1, respectively.

A total of 130 papers was received for consideration from almost 40 different countries. Both the submission and the reviewing process were carried out separately for each workshop even though papers were distributed among reviewers in both program committees according to their particular previously expressed interests.

As a result of this “joint” reviewing process, papers were distributed into three categories: SSPR, SPR, and SSSPR. We designed the technical program of the joint workshop according to the accepted papers and five oral sessions were allocated for each of the two first categories and two oral sessions (eight papers) for the third one.

The two poster sessions (one per workshop) were held at the same time and were allocated a large amount of space and time to encourage discussion and interaction among researchers. A total of 52 papers was selected for oral presentation and 35 papers were presented in the two poster sessions. In addition, we invited five distinguished speakers, Jim Bezdek from the University of West Florida, USA, Marco Gori from the Università di Siena, Italy, Colin de la Higuera from the Université Jean Monnet, Saint Étienne, France, Sarunas Raudys from the Institute of Mathematics and Informatics, Vilnius, Lithuania, and Josef Kittler from the University of Surrey, UK.

SSPR 2000 and SPR 2000 were sponsored by the Conselleria d'Educació de la Generalitat Valenciana under grant RG00-01-22, the Departament de Llenguatges i Sistemes Informàtics (DLSI) and the Escuela Politécnica Superior of the Universitat d'Alacant, the Departament d'Informàtica (DI) of the Universitat de València, and the International Association of Pattern Recognition (IAPR).

We would like to express our gratitude to all our sponsors and, specially, to the members of the two program committees who faced a really tough task which has lead to a selection of papers of a very high quality.

Special thanks are due to Ricardo Ferrís, Esther de Ves, and Elena Díaz of the DI, Universitat de València, and Francisco Moreno-Seco, Jorge Calera-Rubio, and Luisa Micó of the DLSI, Universitat d'Alacant, for their unbeatable effort in the organization of the workshops and in the preparation of the proceedings. We appreciate the help and understanding of the editorial staff of Springer-Verlag, in

A Statistical Matrix Representation Using Sliced Orthogonal Nonlinear Correlations for Pattern Recognition

P. García-Martínez^{1,2}, H.H. Arsenault¹, and C. Ferreira²

¹COPL, Université Laval, Ste-Foy, QC, Canada, G1K 7P4

²Dpt. D'Òptica, Universitat de València, Dr. Moliner, 50. E-46100, Burjassot, Valencia.

SPAIN

Pascuala.garcia@uv.es

1 Introduction

In pattern recognition, the choice of features to be detected is a critical factor to determine the success or failure of a method; much research has gone into finding the best features for particular tasks [1]. When images are detected by digital cameras, they are usually acquired as rectangular arrays of pixels, so the initial features are pixel values. Some methods use those pixel values directly for processing, for instance in normal matched filtering [2], whereas other methods execute some degree of pre-processing, such as binarizing the pixel values [3].

An important tool for pattern recognition is the correlation matrix between objects and its zero-mean cousin, the covariance matrix. Because the curse of dimensionality plagues so many pattern recognition procedures, a variety of methods for dimensionality reduction have been proposed. One of the classical statistical procedures is the principal component analysis [1]. This method (known in the communication theory literature the Karhunen-Loève expansion) finds a lower-dimensional representation that accounts for the variance of the features. The diagonalization of the correlation or covariance matrix is significant for image processing, because among other advantages, it implies the decomposition of images into independent components, it minimizes entropy, it minimizes the mean squared error when some terms are removed, and it is related to principal value decomposition and to factor analysis. Unfortunately the diagonalization of large matrices corresponding to the covariance matrices of images with many pixels is often beyond the capacity of even today's powerful computers.

So it is clear that finding an image decomposition that easily diagonalizes a correlation matrix is of interest; in this paper, we introduce such a decomposition, which we then use in order to see some familiar pattern recognition techniques in a new light, and to propose a new and powerful approach to pattern recognition. The correlation matrix that we propose should not be confused with the classical covariance matrix between objects. Our correlation matrix is between the non-linearly transformed features of two objects.

2 Image Representation in Terms of the SONG Decomposition

Any two-dimensional quantized gray-scale image $f(x, y)$ can be decomposed into a sum of orthogonal elementary images $\{e_m(f)\}$ having the orthogonal property

$$\begin{aligned} e_m(f(x, y))e_n(f(x, y)) &= 0 \quad \text{if } m \neq n \\ e_m(f(x, y))e_n(f(x, y)) &= 1 \quad \text{if } m = n \end{aligned} \quad (1)$$

Each sub-image $\{e_m(f(x, y))\}$ represents a gray level slice of the object. We define the Sliced Orthogonal Nonlinear Generalized (SONG) decomposition of $f(x, y)$ as

$$f(x, y) = \sum_{q=1}^{Q-1} q e_q(f(x, y)) \quad (2)$$

where Q is total number of gray levels in the image and the basis is defined as

$$e_q(f(x, y)) = \begin{cases} 1 & f(x, y) = q \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

A more general definition will be introduced in Section 3. Note that each object point is characterized by only one gray level, so each q -slice is orthogonal to all of the others, as indicated by the orthogonal property. By considering a three-dimensional space with coordinates (x, y, q) , the method can be interpreted as one of placing planes parallel to the (x, y) coordinate plane of the image; each plane then "slices" the function in the area of intersection. All of these areas form disjoint sets of pixels.

The function $e_q(f(x, y))$ is an eigenfunction of the image pixel matrix $f(x, y)$ with eigenvalue q , because $f(x, y)e_q(f(x, y)) = q e_q(f(x, y))$. This property will be important when we consider the correlation matrix.

3 Pattern Recognition in Terms of the SONG Decomposition: The SONG Correlation Matrix

We shall define a SONG correlation [4] based on the new binary SONG decomposition. We shall also relate this correlation to a correlation matrix. Each coefficient of the matrix can be viewed as the cross-correlation between two binary slices of the input scene and of the reference object binarized using the SONG decomposition. The general definition of the SONG correlation matrix allows the representation of common correlation operations as special cases of the correlation matrix.

3.1 Definition of the SONG Correlation

A more general definition of the SONG correlation between an input function $g(x)$ and an object prototype $f(x)$ is

$$\Omega_{gf}(x) = \sum_{q=1}^{Q-1} W(q, \chi) e_q(g(x)) \otimes \sum_{k=1}^{Q-1} W'(k, \chi) e_k(f(x)) \quad (4)$$

where $W(q, \chi)$ and $W'(k, \chi)$ are weighed correlation factors and χ is a parameter that may depend on detection functions or other metrics and \otimes denotes the linear correlation operation. For the sake of clarity, we use one-dimensional functions although we will apply this correlation to two-dimensional images. We introduce a general modified representation of any image as

$$f(x) = \sum_{q=1}^{Q-1} W(q, \chi) e_q(f(x)) \quad (5)$$

Note that if the weight factors are equal to the gray level values, we reconstruct the original function $f(x)$. Then the correlation of Eq. (4) can be expressed in matrix terms as

$$\Omega_{Cross} = \sum_{q=1}^{Q-1} \sum_{k=1}^{Q-1} R_{qk} = \Theta \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1(Q-1)} \\ \vdots & \vdots & \ddots & \vdots \\ R_{(Q-1)1} & R_{(Q-1)2} & \dots & R_{(Q-1)(Q-1)} \end{bmatrix} \quad (6)$$

where $R_{qk} = W(q, \chi) W'(k, \chi) \{e_q(g(x)) \otimes e_k(f(x))\} = W(q, \chi) W'(k, \chi) LC_{gf}^{qk}$. The term LC_{gf}^{qk} represents the linear sub-correlation between the q -slice of $g(x)$ and the k -slice of $f(x)$; the operator Θ means the summation of all the terms of the matrix. It can be viewed as the norm of the matrix, but, the norm of a matrix is defined as $\|\Omega\| = \text{Tr}\{\Omega^T \Omega\}$ where $\text{Tr}\{\cdot\}$ is the trace of a matrix, where the sum of the squared coefficients is used instead of the sum of the amplitudes. So our operator Θ can be viewed as the norm of the matrix in the absolute value sense.

In the following, we shall consider for simplicity objects located such that their correlation peaks appear at the origin (0,0) because the correlation operation is shift invariant, it is trivial to generalize to the case of targets located at any point (x, y) and to the presence of multiple targets, since the correlation is also additive

For the case of autocorrelation where $g(x) = f(x)$, the matrix of Eq. (6) becomes a diagonal matrix because of the orthogonality between the slices, and

$$\Omega_{Auto} = \Theta \begin{bmatrix} A_{11} & \dots & 0 \\ & A_{22} & \vdots \\ \vdots & \vdots & \ddots \\ 0 & 0 & \dots & A_{(Q-1)(Q-1)} \end{bmatrix} \quad (7)$$

The autocorrelation coefficients of the matrix are $A_{qq} = W^2(q, \chi) LC_{gf}^{qq}$. Note that we do not need the squared values of the coefficients because the slices are binary functions.

In the absence of any *a priori* information about the input scene, there is no reason to put any different weights on the binary slices. So the SONG correlation definition that we shall consider is

$$\Omega_{gf}(x) = \sum_{q=1}^{Q-1} [e_q(g(x)) \otimes e_q(f(x))] = \sum_{q=1}^{Q-1} R_{qq} \quad (8)$$

where $W(q, \chi) = W'(k, \chi) = 1$. We are considering only the corresponding q -gray level binary slices for the $g(x)$ and for the $f(x)$ functions. This can be viewed as the summation of the diagonal terms. A similar interpretation is to set to zero all the off-diagonal terms the SONG correlation matrix.

The definition of the SONG correlation of Eq. (8) means that this operation can be viewed as a function that counts the total number of non-zero pixels (or points) in an image at the origin. Indeed, the SONG correlation process consists of separately correlating each binary slice from the image with each binary slice of the prototype corresponding to the same gray level, and then summing the correlation values. Because the base $\{e_q(f)\}$ is orthogonal in the $f(x)$ domain, each correlation value is proportional to the number of pixels that are common to both slices. For the auto-correlation, the sum for all the gray levels yields the total number of pixels in the object. So the SONG auto-correlation at the origin is equivalent to a counting operation: the height of the auto-correlation peak (in the absence of noise) is equal to the number of pixels in the object, and in the case of false targets, the height of the correlation peak is equal to the number of pixels that have the same gray level values in the target and in the prototype. Note that there could be objects that have the same number of pixels but that look totally different from the reference object. But the counting operation that results from our cross-correlation measures the number of equal-valued pixels that are in the same locations for both objects, which is a good measure of similarity.

So coming back to what we pointed out in the introduction, we introduce a diagonal correlation matrix representation for pattern recognition. This can be viewed as a dimensionality reduction of the data, because we used the sum of only the diagonal terms. Moreover, this choice of terms is best because the autocorrelation values (A_{qq} , from Eq. (7)) are found only along the diagonal, where the correlations of the corresponding gray level slices are found. Other off-diagonal terms correspond to the correlations between different gray level slices, and they add a background to the final result without improving the discrimination.

3.2 Common Linear Correlation in Terms of the SONG Matrix

The linear correlation, LC_{gf} , between two quantized gray level functions $g(x)$ and $f(x)$ can be defined in terms of the SONG decomposition as

$$LC_{gf}(x, y) = \left[\sum_{q=1}^{Q-1} q e_q(g(x, y)) \right] \otimes \left[\sum_{k=1}^{Q-1} k e_k(f(x, y)) \right] = \sum_{q=1}^{Q-1} \sum_{k=1}^{Q-1} qk LC_{gf}^{qk} = \Omega_{LC}^{Cross} \quad (9)$$

The subindex indicates that we express the LC as a particular case of the SONG correlation (Ω). So linear correlation can be expressed by means of the SONG matrix using the particular weights as: $W_{LC}^{Auto} = q^2$ and $W_{LC}^{Cross} = qk$. Note that the dimensions of the matrix are $[Q-1 \times Q-1]$. So $(Q-1)^2$ sub-correlations are needed to express the normal linear correlation (which, when done in the usual manner, are all carried out in parallel). The same number of sub-correlations are required for the SONG correlation, and the difference between both correlations are the weight factors. However, because we shall use the particular case of the SONG correlation corresponding to Eq. (8), we need only $(Q-1)$ sub-correlations. The SONG correlation has a significant advantage in discrimination capability because only slices corresponding to the same gray levels are compared.

4 Noise Robustness of the SONG Correlation

One might expect that because the SONG correlation is very selective for object detection and discrimination, it might have poor noise robustness. In this section we consider images that are degraded by very strong Gaussian noise. We prove that whereas other common detection methods like common matched filtering and Phase Only Filtering (POF) are not able to detect the correct object, the SONG correlation will succeed. Moreover, the discrimination capability (DC), measured in terms of one minus the ratio between the cross-correlation and the autocorrelation, will be more stable than for the other methods over a wide range of Gaussian noise levels. The input scene is shown in Fig. 1. It consists of two objects, the reference object being the one in the lower part of the image. This image has 8 gray levels.

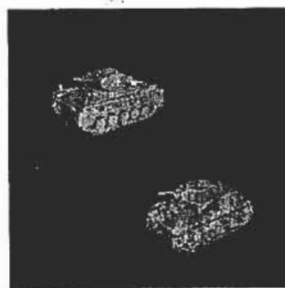


Fig. 1. Input scene with the reference object, placed in the lower part of the figure, and another object to be rejected

Figure 2 is the input scene of Fig. 1 degraded by white additive Gaussian noise ($\sigma=1.9$). The visual pattern information is wiped out by the noise, but as long as some pixels of the image remain unaffected, the SONG correlation will yield a high signal.

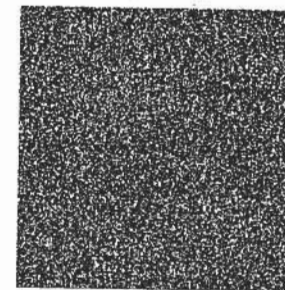


Fig. 2. Input scene highly degraded with Gaussian ($\sigma=1.9$) noise

Table 1 shows the discrimination capability (DC) for the SONG correlation and for other common detection methods. A high value of DC, means that the value of the cross-correlation is low compared to the autocorrelation, which means that a good discrimination and good noise robustness are achieved. On the other hand, a low value of the ratio means that the energy of the cross-correlation has almost the same value that of the auto-correlation.

We generated white Gaussian input noise patterns with various standard deviations σ . Three correlations were considered: the SONG correlation, the linear correlation and the LC using a phase-only filter (POF).

Table 1 shows that for highly degraded images, only the SONG correlation is able to detect the reference object with a high degree of discrimination. Note that the DC for the SONG method is a high stable value for all the noise levels. On the contrary, none of the other methods yield high values of the DC, which implies poor performance for correctly detecting the reference object.

GAUSSIAN NOISE			
Mean=0	SONG Correlation	Linear Correlation	Phase Only Filter
Stand. Dev. (σ)			
0.2	0.95	0.05	0.80
0.25	0.95	0.00	0.70
0.5	0.95	0.00	0.3
0.75	0.95	0.00	0.2
1	0.95	0.00	0.2
1.2	0.95	0.00	0.1
1.9	0.95	0.00	0.1

Table 1. The discrimination capability (DC) of several pattern recognition operations and the new SONG correlation.

5 Conclusion

We have introduced a sliced orthogonal nonlinear generalized (SONG) matrix representation that allows a representation of common linear and nonlinear

correlations. The SONG correlation can be expressed in terms of a sum of linear correlation between the binary gray level slices of an input scene and of a reference object. A weight factor is considered in the definition to allow a more general definition. The discrimination ability and the noise robustness in the presence of Gaussian noise are superior for the SONG method.

Acknowledgements

This work was supported by the Spanish DGES, Dirección General de Enseñanza Superior, project PB96-1134-C02-02 and by an operating grant from the Natural Sciences Engineering Research Council of Canada.

References

- [1] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, (1973). Chapter. 1.
- [2] A. Vander Lugt, A. VanderLugt, " IEEE Trans. Inf. Theory, **IT-10**, 139-145 (1964).
- [3] B. V. K. V. Kumar and D. Casasent, *Appl. Opt.* **20**, 143-147, (1980).
- [4] P. Garcia-Martínez and H. H. Arsenault, *Opt. Commn.* **172**, 181-192 (1999).

A Fast Approximately k -Nearest-Neighbour Search Algorithm for Classification Tasks

Francisco Moreno-Seco, Luisa Micó, and José Oncina *

Dept. Lenguajes y Sistemas Informáticos
Universidad de Alicante, E-03071 Alicante, Spain,
{paco,mico,oncina}@dlsi.ua.es

Abstract. The k -nearest-neighbour (k -NN) search algorithm is widely used in pattern classification tasks. A large set of fast k -NN search algorithms have been developed in order to obtain lower error rates. Most of them are extensions of fast NN search algorithms where the condition of finding exactly the k nearest neighbours is imposed. All these algorithms calculate a number of distances that increases with k . Also, a vector-space representation is usually needed in these algorithms.

If the condition of finding exactly the k nearest neighbours is relaxed, further reductions on the number of distance computations can be obtained. In this work we propose a modification of the LAESA (Linear Approximating and Eliminating Search Algorithm, a fast NN search algorithm for metric spaces) in order to use a certain neighbourhood for lowering error rates and reduce the number of distance computations at the same time.

Keywords: Nearest Neighbour, Metric Spaces, Pattern Recognition.

1 Introduction

Non-parametric classification is one of the most widely used techniques in pattern recognition [2]. One of the simplest techniques (and one of the most popular) is to use the nearest-neighbour (NN) classifier which, given an unknown sample x , finds the prototype p in the training set which is closest to x , then it classifies x in the same class as p . The NN classifier usually obtains acceptable error rates, but it is possible to obtain better (lower) error rates using a number k of nearest neighbours. Thus, a k -nearest-neighbour (k -NN) classifier finds the k nearest neighbours of the sample x , and then, through a voting process, it classifies x in the class which has most representatives among those k nearest neighbours.

Usually, these classifiers are implemented through an exhaustive search; that is, all the distances between the sample and the prototypes in the training set are calculated. When the representation space is an Euclidean space, this exhaustive search is usually not very time-consuming. On the other hand, when working

* The authors wish to thank the Spanish CICYT for partial support of this work through project TIC97-0941.