Lecture Notes in Mathematics 1948

Frédéric Cao · José-Luis Lisani
Jean-Michel Morel · Pablo Musé
Frédéric Sur

# A Theory of Shape Identification

Springer

Frédéric Cao
DxO Labs
3 rue Nationale
92100 Boulogne Billancourt, France
fcao@dxo.com

José-Luis Lisani
Dep. Matemàtiques i Informàtica
University Balearic Islands
ctra. Valldemossa km.7,5
07122 Palma de Mallorca, Balears
Spain
joseluis.lisani@uib.es

Jean-Michel Morel
CMLA, Ecole Normale
Supérieure de Cachan
61 av. du Président Wilson
94235 Cachan Cédex, France
morel@cmla.ens-cachan.fr

Pablo Musé
Instituto de Ingeniería Eléctrica
Facultad de Ingeniería
Julio Herrera y Reissig 565
11300 Montevideo, Uruguay
pmuse@fing.edu.uy

Frédéric Sur
Loria Bat. C - projet Magrit
Campus Scientifique - BP 239
54506 Vandoeuvre-lès-Nancy Cédex
France
sur@loria.fr

ISBN 978-3-540-68480-0 e-ISBN 978-3-540-68481-7
DOI 10.1007/978-3-540-68481-7

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2008927359

Mathematics Subject Classification (2000): 62C05, 62G10, 62G32, 62H11, 62H15, 62H30, 62H35, 68T10, 68T45, 68U10, 91E30, 94A08, 94A13, 94B70



*Cover design*: WMX Design Bender

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

# Preface

Recent years have seen dramatic progress in shape recognition algorithms applied to ever-growing image databases. They have been applied to image stitching, stereo vision, image retrieval, image mosaics, solid object recognition and video and web shape retrieval. More fundamentally, the ability of humans and animals to detect and recognize shapes is one of the enigmas of perception. Digital images and computer vision methods open new ways to address this enigma.

Given a dictionary of digitized shapes and a previously unobserved digital image, the aim of shape recognition algorithms is to know whether some of the shapes in the dictionary are present in the image. This book describes a complete method that starts from a query image and an image database and yields a list of the images in the database containing the query shapes.

Technically speaking there are two main issues. The first is extracting invariant shape descriptors from digital images. Indeed, a shape can be seen from various angles and distances and in various lights. A shape can even be partially occluded by other shapes and still be identifiable. Because the extraction step is so crucial, three acknowledged shape descriptors, SIFT (Scale-Invariant Feature Transform), MSER (Maximally Stable Extremal Regions) and LLD (Level Line Descriptor) will be introduced.[1]

The second issue is deciding whether two shape descriptors are identifiable as the *same shape* or not. This decision process will derive from a unique paradigm, called the Helmholtz principle. For each decision a background model is introduced. Then one decides whether an event of interest (such as the presence of a shape in the image) has occurred if it has a very low probability of occurring by chance in the background model. Thus from the statistical viewpoint shape identification goes back to *multiple hypothesis testing.*

A shape descriptor is recognized if it is not likely to appear by chance in the background model. At a higher complexity level, a group of shape descriptors is recognized if its spatial arrangement could not occur just by chance. These two decisions

[1] In a recent review paper on affine invariant recognition written by a pool of experts, SIFT and MSER were actually acclaimed as the best shape descriptors [122].

rely on simple stochastic geometry and eventually compute a false alarm number for each shape descriptor. The lower this number, the more secure the identification. In that way most familiar simple shapes or images can be reliably identified. Many realistic experiments show false alarm rates ranging from $10^{-5}$ to less than $10^{-300}$.

All in all these lecture notes prove that many shapes can indeed be identified. For these shapes one needs no *a priori* model and no training, just one sample of the shape and what statisticians call a *background model*, or *a null model*. In the case of shape recognition, the term background is to be taken to the letter. By the Helmholtz principle a shape is conspicuous if and only if it cannot be generated by the image background on which it is perceived. The background model is therefore easily learnt from the image database itself.

The above description should not be taken to suggest that the shape recognition problem is solved. The methods described only apply to solid shapes and not to deformable shapes. They only deal with individual shapes and images such as logos or paintings, and not with wide classes of objects such as all humans, all cats or all cars. This latter problem is known as *categorization* and is still widely open to research.

The authors are indebted to their collaborators for many important comments and corrections, particularly to Andrés Almansa, Yann Gousseau and Guoshen Yu. David Mumford and another anonymous referee made valuable comments which reshaped the book. All experiments were done using the public software MegaWave (authors: Jacques Froment and Lionel Moisan). The SIFT method is also public and downloadable.

The present theory was mainly developed at the Centre de Mathématiques et Leurs Applications, at ENS Cachan, at the Universitat de les Illes Balears and at IRISA, Rennes. It was partially financed for the past eight years by the Centre National d'Etudes Spatiales, the Centre National de la Recherche Scientifique, the Office of Naval research (grant N00014-97-1-0839) and the Ministère de la Recherche (project ISII-RNRT), and the Ministerio de Educación y Cultura (project MTM2005-08567). Special thanks to Bernard Rougé and Wen Masters for their great interest and support. We are indebted to Nick Chriss for numerous stylistic corrections.

*Frédéric Cao*
*José Luis Lisani*
*Jean-Michel Morel*
*Pablo Musé*
*Frédéric Sur*

# Contents