

SpringerBriefs in Electrical and Computer Engineering

Signal Processing

Series editors

Woon-Seng Gan, Singapore, Singapore

C.-C. Jay Kuo, Los Angeles, USA

Thomas Fang Zheng, Beijing, China

Mauro Barni, Siena, Italy

More information about this series at <http://www.springer.com/series/11560>

Chen Chen · Yuzhuo Ren · C.-C. Jay Kuo

Big Visual Data Analysis

Scene Classification and Geometric Labeling

 Springer

Chen Chen
Department of Electrical Engineering
University of Southern California
Los Angeles, CA
USA

C.-C. Jay Kuo
University of Southern California
Los Angeles, CA
USA

Yuzhuo Ren
Department of Electrical Engineering
University of Southern California
Los Angeles, CA
USA

ISSN 2191-8112 ISSN 2191-8120 (electronic)
SpringerBriefs in Electrical and Computer Engineering
ISSN 2196-4076 ISSN 2196-4084 (electronic)
SpringerBriefs in Signal Processing
ISBN 978-981-10-0629-6 ISBN 978-981-10-0631-9 (eBook)
DOI 10.1007/978-981-10-0631-9

Library of Congress Control Number: 2016932343

© The Author(s) 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

*Dedicated to my wife and my parents, for
their love and endless support*

—Chen Chen

*Dedicated to my parents for their endless love
and encouragement*

—Yuzhuo Ren

*Dedicated to my wife for her long-term
understanding and support*

—C.-C. Jay Kuo

Preface

Scene understanding is a key issue in computer vision, which recognizes scene image semantic contents and their corresponding contexts. As one of the most challenging scene understanding problems, scene classification considers the semantic concepts in a scene image and classifies scene images into their associated scene categories. Meanwhile, geometric labeling focuses on the scene layouts, where image pixels are labeled and grouped into different geometrically functional classes. It is also an essential step before any further scene understanding tasks such as recognition, annotation, and retrieval, and its results affect the performance of these applications significantly. With rapidly increasing visual data in volume, variety and velocity, traditional approaches toward solutions of the two problems are not adequate in addressing the new challenges. In this brief, we attempt to provide more accurate and scalable solutions for several scene classification and geometric labeling problems: indoor/outdoor classification, outdoor scene classification, and scene layout estimation. This brief will first give an overview on the state-of-the-art algorithms for each problem. Then, it will introduce several techniques for corresponding solutions. They are Expert Decisions Fusion (EDF), Coarse Semantic Segmentation (CSS) and Global-attributes Assisted Labeling (GAL). Extensive experimental results with comparative analysis will be provided for each approaches. Finally, we will conclude and highlight the contributions of these approaches to big visual data analysis.

Contents

1 Introduction	1
References	3
2 Scene Understanding Datasets	7
2.1 Small-Scale Scene Understanding Datasets	7
2.1.1 8-scene Dataset	7
2.1.2 15-scene Dataset	8
2.1.3 UIUC Sports	9
2.1.4 CMU 300	9
2.2 Large-Scale Scene Understanding Datasets	10
2.2.1 80 Million Tiny Image Dataset	11
2.2.2 PASCAL Dataset	12
2.2.3 ImageNet Dataset	13
2.2.4 LabelMe Dataset	13
2.2.5 Scene Understanding (SUN) Dataset	14
2.2.6 Places205 Dataset	19
References	20
3 Indoor/Outdoor Classification with Multiple Experts	23
3.1 Introduction	23
3.2 Individual Indoor/Outdoor Experts	25
3.2.1 Experts from Existing Work	25
3.2.2 Proposed Experts	41
3.3 Data Grouping Using Experts' Decisions	46
3.4 Diversity Gain of Experts Via Decisions Stacking	48
3.4.1 Diversity Gain of Two Experts	49
3.4.2 Construction of Multi-Expert Systems	52
3.5 Expert Decision Fusion Systems	53

3.6	Performance Evaluation	54
3.6.1	Performance of Individual Expert	54
3.6.2	Subspace Classification Performance	57
3.6.3	Scalability	58
3.6.4	Discussion	60
3.7	Summary	61
	References	61
4	Outdoor Scene Classification Using Labeled Segments	65
4.1	Introduction	65
4.2	Review of Previous Works	67
4.2.1	Low-Level Features	67
4.2.2	Mid-Level Features	69
4.2.3	High-Level Features	72
4.2.4	Deep Features	73
4.2.5	Scene Parsing and Semantic Segmentation	75
4.3	Proposed Coarse Semantic Segmentation (CSS)	79
4.3.1	Limitations of Traditional Learning Units	79
4.3.2	Coarse Segmentation	80
4.3.3	Segmental Semantic Labeling	82
4.4	Scene Classification Using CSS	85
4.5	Experimental Results	86
4.5.1	Dataset	86
4.5.2	CSS	86
4.5.3	Scene Classification Results	87
4.6	Summary	89
	References	90
5	Global-Attributes Assisted Outdoor Scene Geometric Labeling	93
5.1	Introduction	93
5.2	Review of Previous Works	94
5.2.1	Geometric Context from a Single Image	94
5.2.2	Blocks World Revisited	95
5.2.3	Single-View 3D Scene Parsing by Attributed Grammar	97
5.2.4	Inferring 3D Layout of Buildings from a Single Image	98
5.3	Proposed GAL System	99
5.3.1	System Overview	99
5.3.2	Initial Pixel Labeling (IPL)	100
5.3.3	Global Attributes Extraction (GAE)	101
5.3.4	Layout Reasoning and Label Refinement (LR2)	108
5.4	Experimental Results	109
5.5	Summary	118
	References	119
6	Conclusion and Future Work	121